

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHQESTXGRSRGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	93.5	2248	2	A35938
2	84	91.3	416	2	A32947
3	75	81.5	591	2	A45135
4	58	63.0	506	1	W2W47
5	55	59.8	822	2	T51049
6	49	53.3	849	2	A96592
7	48	52.2	136	2	T35632
8	47	51.1	471	2	T33997
9	47	51.1	825	2	JC4163
10	47	51.1	836	2	G84727
11	46	50.0	306	2	T21220
12	46	50.0	312	2	A31846
13	46	50.0	399	2	T46259
14	45.5	49.5	135	2	A46398
15	44	47.8	229	2	JC7219
16	44	47.8	843	1	UC719
17	44	47.8	1804	2	T34518
18	44	47.8	2420	2	A84652
19	43.5	47.3	463	2	T51194
20	43	46.7	123	2	T16234
21	43	46.7	141	2	T39554
22	43	46.7	531	2	D86242
23	43	46.7	745	2	G01025
24	43	46.7	774	2	I46609
25	43	46.7	800	2	T02852
26	43	46.7	827	2	T13468
27	43	46.7	842	1	JDVLVS
28	43	46.7	1119	2	T16720
29	43	46.7	1829	2	T35681

30	42.5	46.2	286	2	S50855	neurotrophin-6 - s
31	42	45.7	218	2	S58769	brain-specific pro
32	42	45.7	638	2	I53169	cytokereatin 2 - hu
33	42	45.7	694	2	S71786	wingless receptor
34	42	45.7	778	2	T44761	probable preprotet
35	42	45.7	808	2	F70720	probable transloca
36	42	45.7	1015	2	T42013	frequency clock pr
37	42	45.7	1507	2	B47328	natural killer cel
38	41.5	45.1	553	2	T27245	hypothetical prote
39	41.5	45.1	568	2	H88904	protein Y57G11C.9
40	41.5	45.1	659	2	T27246	hypothetical prote
41	41	44.6	176	2	F71370	probable single-st
42	41	44.6	289	2	AD3394	lipoprotein releas
43	41	44.6	315	2	S53589	SOL2 protein - yea
44	41	44.6	373	2	T02976	probable DNA bindi
45	41	44.6	482	2	C86322	hypothetical prote

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GN>
A:Cross-references: GB:U02929
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: profilaggrin repeat
F:570-893/Region: profilaggrin repeat
F:1074-1397/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 93.5%; Score 86; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 7.3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTXGRSRGRSGSGS 19
DB 551 SHQESTXGRSRGRSGSGS 569

RESULT 2
A32947
filaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
R:McKinnley-grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cammizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A:Title: Characterization of a cDNA clone encoding human filaggrin and localization of ti
A:Reference number: A32947; MUID:89296901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MC>
A:Cross-references: GB:M4355; NID:g182604; PIDN:AA52454.1; PID:g182605
A>Note: The authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:

CjAccession: T35632
Riseejer, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T35632
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-116 <SEB>
A:Cross-references: EMBL:AL079356; PIDN:CAB45627.1; GSPDB:GN00070; SCOEDB:SC669.36C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC669.36C
C:Superfamily: Streptomyces coelicolor probable transposase SC669.36C

Query Match 52.2%; Score 48; DB 2; Length 136;
Best Local Similarity 47.1%; Pred. No. 2.8;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 17
DB 111 AHQHAAGAKKRGRTGA 127

RESULT 8
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
CjSpecies: Caenorhabditis elegans
CjDate: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
CjAccession: T33997
R.Pauley, A.; Scheel, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: EMBL:AF125964; PIDN:AD14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol NZ; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
Best Local Similarity 52.6%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 19
DB 167 SSRSPGRGRGGRSGS 185

RESULT 9
UC4163
DNA-binding protein 5E5 - rat
CjSpecies: Rattus norvegicus (Norway rat)
CjDate: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
CjAccession: UC4163; PC4040
R.Suzuki, E.; Kojima, N.; Yoshimura, K.; Uyemura, K.; Odata, K.; Akagawa, K.
J. Biochem. 118, 122-128, 1995
A:Title: Cloning and sequence analysis of cDNA for a possible DNA-binding protein 5E5 in
A:Reference number: UC4163; WUID:96015159; PMID:8537300
A:Accession: UC4163
A:Molecule type: mRNA
A:Residues: 1-825 <SUZ>
A:Cross-references: DDBJ:D37934; NID:G531260; PIDN:BA07153.1; PID:G531261
A:Experimental source: brain
A:Accession: PC4040
A:Molecule type: protein
A:Residues: 230-455 <SUZ>
C:Comment: This protein has an abundance of arginine, a glycine-rich region and a prolin
C:superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: nerve; phosphoprotein

F436-443/Region: nuclear location signal
F472-731/Region: proline cluster
F462,258,345,360,404,413,570,613,635,752,820/Binding site: phosphate (Ser) (covalent) (b)

Query Match 51.1%; Score 47; DB 2; Length 825;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGRSGS 19
DB 589 GRGRGRSGRSGS 600

RESULT 10
G84727
probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
CjSpecies: Arabidopsis thaliana (mouse-ear cress)
CjDate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
CjAccession: G84727
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Bisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; WUID:20083487; PMID:10617197
A:Accession: G84727
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <STO>
A:Cross-references: GB:AB02093; NID:G4263718; PIDN:AAD15404.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2G32000
A:Map position: 2

Query Match 51.1%; Score 47; DB 2; Length 836;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGRSGS 19
DB 813 GRGRGRSGRSGS 824

RESULT 11
T21220
hypothetical protein F21H7.5 - Caenorhabditis elegans
CjSpecies: Caenorhabditis elegans
CjDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
CjAccession: T21220
R.Gardner, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19393
A:Accession: T21220
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-306 <WIL>
A:Cross-references: EMBL:Z93379; PIDN:CAB07588.1; GSPDB:GN00023; CESP:F21H7.5
A:Experimental source: clone F21H7
C:Genetics:
A:Gene: CESP:F21H7.5
A:Map position: 5
A:Introns: 28/3; 146/3; 232/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 50.0%; Score 46; DB 2; Length 306;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 18
DB 48 SHRSKTKNGSKSGS 65

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comptgen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92
Sequence: 1 SHQESTXGRSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	91.3	416	1	FLIA_HUMAN
2	58	63.0	506	1	VE2_HPV47
3	47	51.1	825	1	5B5_RAT
4	45.5	49.5	135	1	RBP1_DROME
5	44	47.8	843	1	DPO1_HPBVR
6	43	46.7	123	1	LSM4_CAEEL
7	43	46.7	380	1	PEXD_PICPA
8	43	46.7	730	1	DPO1_HPBVA
9	43	46.7	774	1	MRK2_MOUSE
10	43	46.7	842	1	DPO1_HPBVM
11	42	45.7	218	1	P25A_BOVIN
12	42	45.7	638	1	K220_HUMAN
13	42	45.7	694	1	FRZ2_DROME
14	42	45.7	778	1	SEA2_MYCLE
15	42	45.7	808	1	SEA2_MYCTU
16	42	45.7	1015	1	FRQ_CRESP
17	42	45.7	1453	1	NGCR_MOUSE
18	41	44.6	176	1	SSB_TREPA
19	41	44.6	315	1	SOI2_YEAST
20	41	44.6	772	1	MR11_CAEEL
21	41	44.6	777	1	PBBE_VIBCH
22	41	44.6	1337	1	DEXT_STRDO
23	41	44.6	1537	1	ARHB_RAT
24	41	44.6	2233	1	COAC_YEAST
25	41	44.6	2404	1	SON_MOUSE
26	41	44.6	2486	1	SON_HUMAN
27	40.5	44.0	857	1	V2A_CWNT
28	40	43.5	150	1	RL15_PYRHO
29	40	43.5	151	1	RL15_PYRAB
30	40	43.5	308	1	GRP3_ARTSA
31	40	43.5	373	1	YF04_YEAST
32	40	43.5	466	1	CYP8_CAEEL
33	40	43.5	493	1	VE2_HPV19

34	40	43.5	955	1	CLS2_HUMAN	Q9H4D0	homo sapien
35	40	43.5	983	1	Y144_HUMAN	Q14157	homo sapien
36	39.5	42.9	238	1	SFR7_HUMAN	Q16629	homo sapien
37	39.5	42.9	379	1	ROA3_MOUSE	Q8B905	mus musculus
38	39.5	42.9	1115	1	TBC2_CHIRP	Q8VXP3	chlamydomon
39	39.5	42.9	1895	1	WR19_ARATH	Q6Z67	arabidopsis
40	39	42.4	70	1	L2M0_ADE40	Q64858	human adeno
41	39	42.4	168	1	DH1_MAIZE	P12950	zea mays (m
42	39	42.4	255	1	THO4_MOUSE	O08583	mus musculus
43	39	42.4	257	1	THO4_HUMAN	Q66V81	homo sapien
44	39	42.4	419	1	TCF7_MOUSE	O00417	mus musculus
45	39	42.4	484	1	CLK1_HUMAN	P49759	homo sapien

ALIGNMENTS

RESULT 1

ID	FLIA_HUMAN	STANDARD;	PRT;	416 AA.
AC	P20930;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Flilagrin precursor (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88296901; PubMed=2740331;			
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,			
RA	Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.,			
RT	"Characterization of a cDNA clone encoding human flagagrin and			
RT	localization of the gene to chromosome region 1q21."			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).			
RN	[2]			
RP	CITRULLINATION.			
RX	MEDLINE=96374388; PubMed=8780679;			
RA	Senshu T., Kan S., Ogawa H., Marabe M., Asega H.,			
RT	"Preferential determination of keratin K1 and flagagrin during the			
RT	terminal differentiation of human epidermis."			
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).			
CC	-1- FUNCTION: Aggregates keratin intermediate filaments and promotes			
CC	disulfide-bond formation among the intermediate filaments during			
CC	terminal differentiation of mammalian epidermis.			
CC	-1- PTM: Flagagrin is initially synthesized as a large, insoluble,			
CC	highly phosphorylated precursor containing many tandem copies of			
CC	32 AA, which are not separated by "large linker". The precursor			
CC	is deposited as keratohyalin granules. During terminal			
CC	differentiation it is dephosphorylated and proteolytically			
CC	cleaved.			
CC	-1- PTM: Undergoes deamination of some arginine residues			
CC	(citrullination).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M24355; AAA52454.1; -			
DR	PIR; A32947; A32947			
DR	Genew; HGNC:3748; FLG.			
DR	MIM; 135940; -			
DR	GO; GO:0005862; C:intermediate filament; NAS.			
DR	GO; GO:0005198; F:structural molecule activity; NAS.			
DR	GO; GO:0007275; P:development; NAS.			
DR	InterPro; IPR003303; Flilagrin.			

RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandakis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe W., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miles G.L.G.,
 RA Abell J.P., Agbayani A., An H.-O., Andrews-Plambeck C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu B., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatalai B., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Li X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski L.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weltschick G.M., Weissbach U.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye Y., Yen R., Zaveri J.S., Zhan M., Zhang W., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gdbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2195-2195(2000).
 RN [3]

RP SEQUENCE OF 14-52 FROM N.A.
 RX MEDLINE=93109300; PubMed=8417324;
 RA Kim Y.J., Baker B.S.;
 RT "Isolation of RRM-type RNA-binding protein genes and the analysis of
 RT their relatedness by using a numerical approach.";
 RL Mol. Cell. Biol. 13:174-183(1993).
 RT [4]

RP FUNCTION:
 RX MEDLINE=95393975; PubMed=7664738;
 RA Heinrichs V., Baker B.S.;
 RT "The *Drosophila* SR protein RBP1 contributes to the regulation of
 RT doublesex alternative splicing by recognizing RBP1 RNA target
 RT sequences.";
 RL EMBO J. 14:3987-4000(1995).

RP FUNCTION: Contributes to the activation of female-specific DSX
 CC splicing in vivo by recognizing the RBP1 target sequences within
 CC the purine-rich polyrimidine tract of the female-specific 3'
 CC splice site.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RBPl-A;
 CC IsoId=Q02427-1; Sequence=Displayed;
 CC Name=2; Synonyms=RBPl-B;
 CC IsoId=Q02427-2; Sequence=VSP_005817;
 CC -1- TISSUE SPECIFICITY: Ubiquitous;
 CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
 CC -1- PTM: Extensively phosphorylated on serine residues in the RS
 CC domain (Probable).
 CC -1- SIMILARITY: Belongs to the splicing factor SR family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; L04929; AAA28850.1; -.
CC DR EMBL; AE003668; AAF54555.1; -.
CC DR EMBL; S51691; AAB24622.1; -.
CC DR PIR; A46398; A46398.
CC DR PIR; A48110; A48110.
CC DR FlyBase; FBgn010252; Rbp1.
CC DR InterPro; IPR000504; RNA_rec_mot.
CC DR Pfam; PF00076; xrm; 1.
CC DR PROSITE; PS50102; RMY; 1.
CC DR PROSITE; PS00303; RMY_RNP_1; FALSE_NEG.
CC KW Nuclear protein; RNA-binding; mRNA splicing; alternative splicing;
CC phosphorylation.
CC KM
CC FT DOMAIN 11 84 RNA-BINDING (RRM).
CC FT FT 82 135 ARG/SER-RICH (RS DOMAIN).
CC FT FT VASAPLIC 107 135 Missing (in isoform 2).
CC FT FT /FtId=VSP_005817.
CC FT FT CONFLICT 14 14 Y -> F (IN REF. 3).
CC FT FT CONFLICT 65 65 R -> A (IN REF. 1).
CC SQ SEQUENCE 135 AA; 15446 MW; 896DCB902518D991 CRC64;
CC -----
CC Query Match 49.5%; Score 45.5; DB 1; Length 135;
CC Best Local Similarity 52.2%; Pred.No. 1.7;
CC Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1
CC -----
Cc 4 ESTXGRSR-----GRSGRGS 19
Cc | : ||||| | |||||
Cc Db 80 EMSSGGRDRRRRGSGSSGRSGS 102
Cc -----
Cc RESULT 5
Cc DPOL_HPBVR
Cc ID_DPOL_HPBVR STANDARD; PRT; 843 AA.
Cc AC P03157;
Cc DT 21-JUL-1986 (Rel. 01, Created)
Cc DT 21-JUL-1986 (Rel. 01, Last sequence update)
Cc DT 16-OCT-2001 (Rel. 40, Last annotation update)
Cc DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
Cc directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
Cc OS Hepatitis B virus (subtype adt.).
Cc OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
Cc OX NCBI_TaxID=106820;
Cc RN [1]
Cc RP SEQUENCE FROM N.A.
Cc RA MEDLINE=8316819; PubMed=6300776;
Cc RX Ono Y., Onoda H., Saeeda R., Igarashi K., Sugino Y., Nishioka K.;
Cc RT "The complete nucleotide sequences of the cloned hepatitis B virus
Cc DNA; subtype adr and adw.";
Cc RL Nucleic Acids Res. 11:1747-1757(1983).
Cc CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
Cc + [DNA] (N).
Cc CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
Cc phosphomonoester.
Cc CC -----
Cc CC This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc DR EMBL; V00867; -, NOT ANNOTATED_CDS.
Cc PIR; A00704; JDLVLR.

```

```

DR InterPro: IPR001462; DNAPOL_viral_C.
DR InterPro: IPR000201; DNAPOL_viral_N.
DR InterPro: IPR000477; RVase.
DR Pfam: PF00336; DNA_pol_viral_C; 1.
DR Pfam: PF00242; DNA_pol_viral_N; 1.
DR Pfam: PF00078; rvt; 2.
DR ProDom: PD000814; DNAPOL_viral_C; 1.
DR TransErase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KM Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C4B8B CRC64;

Query Match 47.8%; Score 44; DB 1; Length 843;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGRSGRSGSGS 19
DB 226 QPCGGSVARGRSGSGS 242

RESULT 6
LSM4_CAEEL STANDARD; PRT; 123 AA.
AC Q19952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.
GN F32A5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pseudodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Pauley A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases
CC -! FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -! SUBUNIT: LSM subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -! SUBCELLULAR LOCATION: Nuclear (potential).
CC -! SIMILARITY: Belongs to the snRNP Sm proteins family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20864; AAC46661.1; -.
DR PIR: T16234; T16234.
DR WormPep: F32A5.7; CE01277.
DR InterPro: IPR006649; snRNP.
DR InterPro: IPR001163; snRNP_Sm.
DR Pfam: PF01423; LSM; 1.
DR ProDom: PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
DR Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
KM RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 3.8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGRSGRSGSGS 18
DB 93 REGSRGRGRGRGRG 108

```

```

RESULT 7
ID_PEXD_PICPA STANDARD; PRT; 380 AA.
AC Q92266;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peroxisomal membrane protein PEX13 (Peroxin-13).
GN PEX13.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4922;
[1]
RN SEQUENCE FROM N.A., AND MUTAGENESIS.
RP MEDLINE=9701155; PubMed=8958165;
RA Gould S.J., Kalish J.E., Morrell J.C., Bjorkman J., Urquhart A.J.,
RA Crane D.I.;
RT Factor for the predominantly cytoplasmic Pst1 receptor.
RT J. Cell Biol. 135:85-95(1996).
CC -! FUNCTION: Component of the peroxisomal translocation machinery
CC with PEX14 and PEX17. Interacts with the Pst1 and Pst2 proteins.
CC (Pst1/PEX5). Involved in the import of Pst1 and Pst2 proteins.
CC -! SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -! ALTERNATIVE PRODUCTS:
CC Event-Alternative initiation;
CC Comment=2 isoforms, PEX13L (shown here) and PEX13S, are produced
CC by alternative initiation;
CC -! SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U70067; AAB09087.1; -.
DR HSP: Q64010; ICKA.
DR InterPro: IPR007223; Peroxin-13_N.
DR InterPro: IPR001452; SH3.
DR Pfam: PF04088; Peroxin-13_N; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD00066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KM Peroxisome; Transport; Protein transport; Transmembrane; SH3 domain;
KM Alternative initiation.
SQ CHAIN 1 380
FT CHAIN 33 380
FT INIT MET 33 33
FT DOMAIN 1 230
FT TRANSMEM 231 251
FT DOMAIN 252 380
FT DOMAIN 277 344
FT MUTAGEN 286 286
FT MUTAGEN 287 287
FT MUTAGEN 288 288
FT MUTAGEN 291 291
FT MUTAGEN 296 296
SQ SEQUENCE 380 AA; 40695 MW; EBBAB39F93BA832 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 380;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 SHQESTXGRSGRSGSGS 18

```

```

Db          315 SNOESTWKCRRDQKVG 332
      ||||| : ||| :
RESULT 8
DPOL_HPBV4
ID_DPOL_HPBV4      STANDARD;      PRT;      730 AA.
AC      p12933;
DT      01-OCT-1989 (Rel. 12, Created)
DT      01-OCT-1989 (Rel. 12, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE      directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN      P.
OS      Hepatitis B virus (subtype adr4).
OC      Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX      NCBI_TaxID=10409;
RN      SEQUENCE FROM N.A.
RP      MEDLINE=83246570; Pubmed=6306594;
RA      Fujiyama A., Miyahara A., Nozaki C., Yoneyama T., Ohnomo N.,
RA      Matsubara K.,
RT      Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT      adr.
RL      Nucleic Acids Res. 11:4601-4610(1983).
CC      -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC      + {DNA}(N).
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC      phosphonooxester.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X01587; CA25746.1; -
DR      InterPro; IPR001462; DNADPOL_viral_C.
DR      InterPro; IPR000201; DNADPOL_viral_N.
DR      InterPro; IPR000477; RVTse.
DR      Pfam; PF00336; DNA_pol_viral_C_1.
DR      Pfam; PF00242; DNA_pol_viral_N_1.
DR      Pfam; PF00078; tvc; 2.
DR      ProDom; PD000814; DNADPOL_viral_C_1.
KW      Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW      Hydroxylase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ      SEQUENCE 730 AA; 82207 MW; 898A8BB50D496F2 CRC64;

Query Match          46.7%; Score 43; DB 1; Length 730;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      10 SRGSRGSGS 19
      |||||
Db          233 ARGSRGSGS 242

```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Embryo;
RX      MEDLINE=93364122; Pubmed=8358177;
RA      Inglis J.D., Lee M., Hill R.E.;
RT      "Emk", a protein kinase with homologs in yeast maps to mouse
RT      chromosome 19."
RL      Mamm. Genome 4:401-403(1993).
CC      -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC      -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC      MARK subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X70764; CA50040.1; -.
DR      PIR; I48609; I48609.
DR      HSSP; Q63450; I406.
DR      MGD; MGI:99638; Mark2.
DR      InterPro; IPR001772; Kinase Cterm.
DR      InterPro; IPR000719; Prot Kinase.
DR      InterPro; IPR008271; Ser Thr pkin AS.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR001245; Tyr_kinase.
DR      Pfam; PF02149; KAI; 1.
DR      Pfam; PF00069; pkinase; 1.
DR      PRINTS; PR00109; TYRKINASE.
DR      ProDom; PD000001; Prot Kinase; 1.
DR      SMART; SM00220; S_TKC_1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST_1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
KW      Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT      DOMAIN 53 304
FT      NP_BIND 59 67
FT      BINDING 82 82
FT      ACT_SITE 175 175
FT      ACT_SITE 175 175
SQ      SEQUENCE 774 AA; 85874 MW; 02BFD7B5443493A CRC64;

Query Match          46.7%; Score 43; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      5 STXGSRGSRGSGS 19
      |||||
Db          614 SPSGHSQGRGSGS 628
      |||||
RESULT 10
DPOL_HPBVM
ID_DPOL_HPBVM      STANDARD;      PRT;      842 AA.
AC      P31870; Q67861; Q67866;
DT      01-JUN-1993 (Rel. 26, Created)
DT      01-JUN-1993 (Rel. 26, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE      directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN      P.
OS      Hepatitis B virus (subtype adr / mutant).
OC      Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX      NCBI_TaxID=31512;
RN      [1]
RP      SEQUENCE FROM N.A.
RP      MEDLINE=89183619; Pubmed=2928116;
RA      Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RA      "The nucleotide sequence and reading frames of a mutant hepatitis B
RT      virus subtype adr."
RL      Nucleic Acids Res. 17:2124-2124(1989).

```

```

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5' -
CC phosphomonester.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X14193; CA33299.1; -
CC EMBL; X14193; CA332405.1; ALT_TERM.
CC PIR; S04568; JDLV6.
CC InterPro; IPR001462; DNAPOL_viral_C.
CC InterPro; IPR002021; DNAPOL_viral_N.
CC InterPro; IPR000477; RVTse.
CC Pfam; PF00336; DNA_pol_viral_C; 1.
CC Pfam; PF00242; DNA_pol_viral_N; 1.
CC Pfam; PF00078; Ivt; 2.
CC ProDom; PD000814; DNAPOL_viral_C; 1.
CC Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
CC Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
CC SEQUENCE 842 AA; 94545 MW; 137FE37246BA5CC6 CRC64;
-----
Query Match 46.7%; Score 43; DB 1; Length 842;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
-----
QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242
-----
RESULT 11
ID P25A_BOVIN STANDARD; PRT; 218 AA.
AC Q2957; O9TRW3; O9TRW4;
DT 01-NOV-1987 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 25 kDa brain-specific protein (p25-alpha).
OS Bos taurus (Bovine).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC Bovidae; Bovinae; Bos.
CC NCBI_TaxID=9913;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=Brain;
CC MEDLINE=95375016; PubMed=7647094;
CC "A novel brain-specific 25 kDa protein (p25) is phosphorylated by a
CC Ser/Thr-Pro kinase (TPK II) from tau protein kinase fractions.";
CC FBS Lett. 289:37-43(1991).
CC -1- TISSUE SPECIFICITY: Brain specific.
CC -1- PTM: Phosphorylated by TPK II.
CC -1- SIMILARITY: Belongs to the p25 family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X85738; CA59741.1; -
CC PIR; S58769; S58769.
CC InterPro; IPR008907; P25-alpha.
CC Pfam; PF05517; P25-alpha; 1.
CC Phosphorylation.
CC SEQUENCE 218 AA; 23472 MW; 55F5DAB42DC3A638 CRC64;
-----
Query Match 45.7%; Score 42; DB 1; Length 218;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 9; Conservative 3; Mismatches 4; Indels 2; Gaps 1;
-----
QY 1 SHOR-STXGSRGSGR 16
DB 174 SHKRPDPGSGRGRAGR 191
-----
RESULT 12
ID K220_HUMAN STANDARD; PRT; 638 AA.
AC Q01546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2p) (K2p)
DE (CK 2p).
GN KR12P.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=93114504; PubMed=1282112;
CC Collin C., Ohnayeun J.P., Grund C., Franke W.W.;
CC "Suprabasal marker proteins distinguishing keratinizing squamous
CC epithelia: cyokeratin 2 polypeptides of oral masticatory epithelium
CC and epidermis are different."
CC Differentiation 51:137-148(1992)
CC -1- FUNCTION: Probably contributes to terminal cornification.
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC keratinocytes.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC (neutral to basic; 56-70 kDa) [K1 to K8]
CC -1- SIMILARITY: Belongs to the intermediate filament family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M99063; AAA3746.1; -
CC PIR; I53169; I53169.
CC GO; GO:0005882; C:intermediate filament; NAS.
CC GO; GO:0005198; F:structural molecule activity; NAS.
CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
CC InterPro; IPR001664; IF.
CC InterPro; IPR002957; Keratin_I.
CC InterPro; IPR003054; Keratin_II.
CC Pfam; PF00038; filament; 1.
CC PRINTS; PR01246; TYPE1KERATIN.
CC PRINTS; PR01276; TYPE2KERATIN.
CC PROSITE; PS00226; IF_1.
CC Intermediate filament; Coiled coil; Keratin; Phosphorylation.

```



```

FT DOMAIN 1 182 HEAD.
FT DOMAIN 183 492 ROD.
FT DOMAIN 493 638 TAIL.
FT DOMAIN 183 218 COIL 1A.
FT DOMAIN 219 237 LINKER 1.
FT DOMAIN 238 329 COIL 1B.
FT DOMAIN 330 353 LINKER 12.
FT DOMAIN 354 492 COIL 2.
FT MOD RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 638 AA; 65871 MW; 9B743AB872076AF CRC64;

Query Match 45.7%; Score 42; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGRS 17
   :|:|:|:|:|:|
Db 7 KKSFSGRSGRSGRS 21

RESULT 13
FR22 DROME STANDARD; PRT; 694 AA.
AC 09VXX3; Q94916; Q9VXX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fizzled protein 2 precursor (Fizzled-2) (dfz2).
GN FZ2 OR CG9739/CG14083.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;

[1]
RN SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
RX MEDLINE=96353971; PubMed=8717036;
RA Bharat P., Brink M., Samos C.H., Heish J.C., Wang Y., Macke J.P.,
RA Andrew D., Nathans J., Nusse R.;
RT "A new member of the frizzled family from Drosophila functions as a
RT WINGLESS receptor.";
RL Nature 382:225-230(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celtniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Achyavani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Beeson R.M., Basu A.A., Baxendale U., Bayraktiroglu U., Beasley E.M.,
RA Borrova D., Bonos P.V., Berman B.F., Bhattacharya D., Bolshakov S.,
RA Chertis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Churry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.B., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kinnison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mantei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Modyarty C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacle J.M.,
RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
CC are coupled to the beta-catenin canonical signaling pathway, which
CC leads to the activation of dishevelled proteins, inhibition of
CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
CC of Wnt target genes. A second signaling pathway involving PKC and
CC calcium fluxes has been seen for some family members, but it is
CC not yet clear if it represents a distinct pathway or if it can be
CC integrated in the canonical pathway, as PKC seems to be required
CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
CC to involve interactions with G-proteins. Required to coordinate
CC the cytoskeletons of epidermal cells to produce a parallel array
CC of cuticular hairs and bristles.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).

CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
CC between 15 and 70 per cent of egg length, including the
CC invaginating cells of the ventral furrow. Stripe pattern is
CC emerging by early stage 8. From stage 9 and continuing throughout
CC embryogenesis, expression is seen in the developing CNS. At stage
CC 10, expressed in 15 stripes in the presumptive head and trunk
CC regions, in the posterior midgut primordium, in a subset of cells
CC of anterior midgut invagination and in the procephalic lobe. At
CC stage 12, expression declines in epidermis and increases in the
CC midgut and visceral mesoderm. At stage 17, only expressed in the
CC CNS, hindgut and dorsal vessel.

CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
CC the Wnt/beta-catenin signaling pathway (By similarity).

CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.

CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
CC family.

CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U65589; AAC47273.1; -
DR EMBL: AB003518; AAF49185.2; -
DR PIR: S71786; S71786.
DR FLYBASE: FBgn0016797; Fz2.
DR GO: GO:0016021; C:integral to membrane; NMS.
DR GO: GO:0017147; F:Wnt-protein binding; IDA.
DR GO: GO:0006928; P:cell motility; IMP.
DR GO: GO:0007163; P:establishment and/or maintenance of cell po. .; ISS.
DR GO: GO:0008585; P:female gonad development; IMP.
DR GO: GO:0016055; P:Wnt receptor signaling pathway; IDA.
DR InterPro: IPR000539; Fizzled.
DR InterPro: IPR000024; Fz domain.
DR InterPro: IPR000832; GPCR secretin.
DR Pfam: PF01534; Fizzled.1.
DR Pfam: PF01392; Fz.1.
DR PRINTS: PR00489; FRIZZLED.
DR SMART: SM00063; FRI.1.
DR PROSITE: PS50038; FZ.1.
DR PROSITE: PS50261; G-PROTEIN RECEPTOR_FZ_4.1.
DR MultiGene Family: Receptor; G-protein coupled receptor; Transmembrane;
DR Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 694 FRIZZLED PROTEIN 2.

```

```

FT DOMAIN 23 315 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 316 336 1 (POTENTIAL).
FT DOMAIN 337 352 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 353 373 2 (POTENTIAL).
FT DOMAIN 374 397 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 398 418 3 (POTENTIAL).
FT DOMAIN 419 439 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 440 460 4 (POTENTIAL).
FT DOMAIN 461 482 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 483 503 5 (POTENTIAL).
FT DOMAIN 504 534 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 535 555 6 (POTENTIAL).
FT DOMAIN 556 584 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 585 605 7 (POTENTIAL).
FT DOMAIN 606 624 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 625 644 GLY-RICH.
FT DOMAIN 645 664 F2.
FT TRANSSEM 665 684 LYS-THR-X-X-TRP MOTIF.
FT DOMAIN 685 704 PDZ-BINDING.
FT TRANSSEM 705 724 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 725 744 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TRANSSEM 745 764 V -> A (IN REF. 1).
FT CONFLICT 765 784 S -> T (IN REF. 1).
FT TRANSSEM 785 804 S -> T (IN REF. 1).
SQ SEQUENCE 694 AA; 75451 MW; 6C510F13BAFB096 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 694;
Best Local Similarity 52.6%; Pred. No. 34;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

CY 1 SHQSTYGRSGRSGRSGS 19
Db 182 SYTAGSGGSGSGSGSGS 200

```

```

RESULT 14
SEQ2 MYCTU STANDARD; PRT; 778 AA.
ID SEQ2 MYCTU STANDARD; PRT; 778 AA.
AC 032922;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Preprotein translocase secA 2 subunit.
OS SEC2 OR SEC2 OR M2082 OR MLCB1788.45C.
CN Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=TN.
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Jancic C., Maclean J., Mole S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RT Nature 409:1007-1011 (2001).
RL
CC -i- FUNCTION: Involved in protein export. Interacts with the secY/secE
subunits. SecA has a central role in coupling the hydrolysis of
ATP to the transfer of pre-secretory periplasmic and outer
membrane proteins across the membrane (By similarity).
CC -i- SUBUNIT: Part of the prokaryotic protein translocation apparatus
which comprises secA, secB, secE, secF, secH and secY (By
similarity).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (By
similarity).
CC -i- SIMILARITY: Belongs to the secA family.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC DR EMBL; AL008609; CAA15477.1; -
CC DR EMBL; AL583924; CAC1037.1; -
CC DR PIR; T44761; T44761.
CC DR Leproma; ML2082; -.
CC DR InterPro; IPR001650; Helicase_C.
CC DR InterPro; IPR00185; SecA.
CC DR Pfam; PF00271; helicase_C/1.
CC DR Pfam; PF01043; SecA protein; 1.
CC DR PRINTS; PR00306; SecA.
CC DR PROSITE; PS01312; SEC2A.1.
CC DR Protein transport; ATP-binding, Membrane; Translocation; Transport;
CC NP BIND 109 116 ATP (POTENTIAL).
FT SEQUENCE 778 AA; 84814 MW; E855CDD2C023A1E1 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 778;
Best Local Similarity 47.1%; Pred. No. 38;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

CY 2 HQESTYGRSGRSGRSG 18
Db 530 HTERLDNQLRGRAGRG 546

```

```

RESULT 15
SEQ2 MYCTU STANDARD; PRT; 808 AA.
ID SEQ2 MYCTU STANDARD; PRT; 808 AA.
AC 050612;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Preprotein translocase secA 2 subunit.
OS SEC2 OR RV1821 OR M1869 OR MTCY1A11.22C OR MB1852.
CN Mycobacterium tuberculosis.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
[1]
RN SEQUENCE FROM N.A.
RC SPECIES=M. tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagsels K., Krogh A., McLean J., Mole S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulton J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RT Nature 393:537-544 (1998).
RL
CC -i- SEQUENCE FROM N.A.
CC SPECIES=M. tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=1218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Dittelback T., Weidman J., Kouti H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RT J. Bacteriol. 184:5479-5490 (2002).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: Involved in protein export. Interacts with the secY/secE
CC subunits. SecA has a central role in coupling the hydrolysis of
CC ATP to the transfer of pre-secretory periplasmic and outer
CC membrane proteins across the membrane (By similarity).
CC -1- SUBUNIT: Part of the prokaryotic protein translocation apparatus
CC which comprise secA, secB, secD, secE, secF, secG and secY
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (By
CC similarity).
CC -1- SIMILARITY: Belongs to the secA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z78020; CAB01458.1; -.
DR EMBL: AE007045; AAK46142.1; -.
DR EMBL: BX248340; CAD94555.1; -.
DR PIR: F70720; F70720.
DR TIGR: MT1869; -.
DR TubercuList; RV1821; -.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR000185; SecA.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF01643; SecA protein; 1.
DR PRINTS: PR00906; SEC_A.
DR PROSITE: PS01312; SEC_A; 1.
DR Protein transport; ATP-binding; Membrane; Translocation; Transport;
KW Complete proteome.
KW NP BIND 139 146 ATP (POTENTIAL).
SQ SEQUENCE 808 AA; 88951 MW; 6AF6707CC3951D2F CRC64;

```

```

Query Match 45.7%; Score 42; DB 1; Length 808;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 HQESTYGRSGRSGRSG 18
DB 564 HTERLDNQLRGRGRG 580

```

Search completed: September 28, 2004, 06:05:45
 Job time : 10.224 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds

(without alignments)
149,211 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHQESTYGRSGRGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rvlinus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	97.8	797	4	Q16824	Q16824 homo sapien
2	86	93.5	990	4	Q15206	Q15206 homo sapien
3	86	93.5	1218	4	Q05331	Q05331 homo sapien
4	84	91.3	798	4	Q9H4U3	Q9H4U3 homo sapien
5	84	91.3	1084	4	Q01212	Q01212 homo sapien
6	75	81.5	591	4	Q01720	Q01720 homo sapien
7	75	81.5	687	4	Q9H4U2	Q9H4U2 homo sapien
8	70	76.1	465	4	Q03838	Q03838 homo sapien
9	69	75.0	322	4	Q05370	Q05370 homo sapien
10	55	59.8	822	3	Q9P312	Q9P312 neurospira
11	50	54.3	1284	5	Q9V8Y3	Q9V8Y3 drosophila
12	49	53.3	820	12	Q6R375	Q6R375 mouse cytom
13	49	53.3	849	10	Q9C7M2	Q9C7M2 arabidopsis
14	49	53.3	849	10	Q949M6	Q949M6 arabidopsis
15	49	53.3	1142	13	Q7S2V2	Q7S2V2 xenopus lae
16	48	52.2	481	16	Q7UUP9	Q7UUP9 rhodospirillum rubrum

17	48	52.2	1229	3	Q86ZH1	Q86ZH1 neurospira
18	47	51.1	471	5	Q9AYO	Q9AYO caenorhabditis
19	47	51.1	836	10	Q9SKZ9	Q9SKZ9 arabidopsis
20	46	50.0	306	5	Q45386	Q45386 caenorhabditis
21	46	50.0	399	4	Q9NTA9	Q9NTA9 homo sapien
22	46	50.0	436	16	Q82C67	Q82C67 streptococcus
23	46	50.0	467	4	Q9BW18	Q9BW18 homo sapien
24	46	50.0	700	13	Q42378	Q42378 brachydanio
25	46	50.0	715	13	Q8CRU0	Q8CRU0 brachydanio
26	46	50.0	715	13	Q8CGS8	Q8CGS8 brachydanio
27	46	50.0	737	13	Q42107	Q42107 brachydanio
28	46	50.0	737	4	Q9BQ39	Q9BQ39 homo sapien
29	46	50.0	787	5	Q9GSD4	Q9GSD4 plasmodium
30	46	50.0	843	12	Q9QNS2	Q9QNS2 hepatitis b
31	45.5	49.5	144	5	Q8T9K6	Q8T9K6 drosophila
32	45	48.9	96	15	Q8JRU5	Q8JRU5 human immun
33	45	48.9	96	15	Q9WRE7	Q9WRE7 human immun
34	45	48.9	96	15	Q9WRE4	Q9WRE4 human immun
35	45	48.9	96	15	Q8JRV4	Q8JRV4 human immun
36	45	48.9	96	15	Q8ADP5	Q8ADP5 human immun
37	45	48.9	243	10	Q9AX45	Q9AX45 oryza sativa
38	45	48.9	247	10	Q8GV23	Q8GV23 chlamydomonas
39	45	48.9	426	4	Q8WV76	Q8WV76 homo sapien
40	45	48.9	476	12	Q80890	Q80890 herpesvirus
41	45	48.9	652	5	Q77158	Q77158 apis mellifera
42	44.5	48.4	319	10	Q94AJ0	Q94AJ0 arabidopsis
43	44	47.8	96	15	Q8ADX3	Q8ADX3 human immun
44	44	47.8	229	11	Q9JW93	Q9JW93 mus musculus
45	44	47.8	259	4	Q86YA2	Q86YA2 homo sapien

ALIGNMENTS

RESULT 1						
ID	Q16824	PRELIMINARY;	PRT;	797 AA.		
AC	Q16824					
DT	01-NOV-1996	(TRENBLREL. 01, Created)				
DT	01-NOV-1996	(TRENBLREL. 01, Last sequence update)				
DT	01-UN-2003	(TRENBLREL. 24, Last annotation update)				
DE	Profilaggrin (Fragment).					
GN	FLG.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=91064347; PubMed=2248957;					
RA	Gan S.Q., McBride C.W., Idler W.W., Markova N., Steinhert P.M.;					
RT	"Organization, structure, and polymorphisms of the human profilaggrin					
RT	gene [published erratum appears in Biochemistry 1991 Jun					
RT	11;30(23):5814.];					
RL	Biochemistry 29:9432-9440(1990).					
DR	EMBL; M60502; AAA63248.1; -					
DR	GO; GO:0005198; F:structural molecule activity; IEA.					
DR	InterPro; IPR003303; Filaggrin.					
DR	Pfam; PF03515; Filaggrin.4.					
DR	PRINTS; PR00487; FILAGGRIN.					
FT	NON_TER					
FT	SEQUENCE 797 AA; 85176 MW; 60E6184763BDA6B CRC64;					
Query Match	97.8%; Score 90; DB 4; Length 797;					
Best Local Similarity	94.7%; Pred. No. 5.8e-06;					
Matches	18; Conservative 0; Mismatches 1; Indels 0;					
QY	1 SHQESTYGRSGRGRSGS 19					
DB	427 SHQESTYGRSGRGRSGS 445					
RESULT 2						

Q15206 ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene [published erratum appears in Biochemistry 1991 Jun 11;30(23):5814]."
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene".
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1; -;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
Query Match 93.5%; Score 86; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 3.2e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SHQSTXGRSGRSGSGS 19
Db 551 SHQSTXGRSGRSGSGS 569
RESULT 3
Q05331 ID Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein."
RL Mol. Cell. Biol. 13:613-625(1993).
CC -I- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -I- PM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
CC -I- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1; -;
DR PIR; A48118; A48118.
DR HSSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_5100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00303; S100_CABP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;
Query Match 93.5%; Score 86; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SHQSTXGRSGRSGSGS 19
Db 773 SHQSTXGRSGRSGSGS 791
RESULT 4
Q094U3 ID Q094U3 PRELIMINARY; PRT; 798 AA.
AC Q094U3;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE D1AN1.1.2 (Profilaggrin 3' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER 798
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 5.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 SHQSTXGRSGRSGSGS 19
Db 428 SHQSTXGRSGRSGSGS 446
RESULT 5
Q01212 ID Q01212 PRELIMINARY; PRT; 1084 AA.

AC 001212; 003840;
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinhart P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]."
 RT 11;30(23):5814.
 RL Biochemistry 29:9432-9440(1990).
 DR EMBL; M60503; AAA63243.1; -.
 DR EMBL; M60501; AAA63243.1; JOINED.
 DR GO; GO:0005882; C:intermediate filament; NAS.
 DR GO; GO:0005198; F:structural molecule activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PR00487; Filaggrin.
 DR NON TER 1
 FT SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;
 SQ

Query Match 91.3%; Score 84; DB 4; Length 1084;
 Best Local Similarity 89.5%; Pred. No. 7.3e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 Db 64 SHOESTXGRSGRSGSGS 82

RESULT 6
 Q001720
 ID 001720; PRELIMINARY; PRT; 591 AA.
 AC 001720;
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
 DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Presland R.B., Haydock P.V., Fleckman P., Natusneksiri W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 organization and identification of an S-100-like calcium binding
 domain at the amino terminus."
 RT domain at the amino terminus."
 RL J. Biol. Chem. 267:23772-23781(1992).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEN COPIES OF
 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS
 DR EMBL; L01089; AAA60177.1; -.
 DR EMBL; L01090; AAA60176.1; -.
 DR

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; Filaggrin.
 DR PROSITE; PS00018; EF HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Polymorphism.
 FT PROPEP 1 293
 FT CHAIN 294 467
 FT PROPEP 468 474
 FT CHAIN 475 >591
 FT CA BIND 19 32
 FT CA BIND 62 73
 FT NON TER 591 591
 FT SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;
 SQ

Query Match 81.5%; Score 75; DB 4; Length 591;
 Best Local Similarity 84.2%; Pred. No. 0.0011;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 Db 449 SHOESTXGRSGRSGSGS 467

RESULT 7
 Q9H4U2
 ID Q9H4U2; PRELIMINARY; PRT; 687 AA.
 AC Q9H4U2;
 DT 01-MAR-2001 (TREMELREL. 16, Created)
 DT 01-MAR-2001 (TREMELREL. 16, Last sequence update)
 DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
 DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL356504; CAC13172.1; -.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; Filaggrin.
 DR PROSITE; PS00018; EF HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON TER 687 687
 FT SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;
 SQ

Query Match 81.5%; Score 75; DB 4; Length 687;
 Best Local Similarity 84.2%; Pred. No. 0.0012;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY      1 SHOESTXGRSGRSGSGS 19
      ||||| |||||
DB      449 SHOESTXGRSGRSGSGS 467

RESULT 8
ID      003838      PRELIMINARY;      PRT;      465 AA.
AC      003838:
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT      01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE      FILAGGRIN (PROFILAGGRIN) (Fragment).
GN      FILG.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RX      MEDLINE=91064347; PubMed=2248957;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 29:9432-9440(1990).
RN      (2)
RP      REVISIONS.
RX      MEDLINE=91255199; PubMed=2043621;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 30:5814-5814(1991).
CC      -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
      PROMOTES DILIPID-BOND FORMATION AMONGST THE INTERMEDIATE
      FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
      -1- POLYMORPHISMS: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
      WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
      REPEATS.
      -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
      INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEN
      COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN
      GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
      PROTEOLYTICALLY CLEAVED.
      EMBL: M60499; AAA63246.1; -.
      DR      GO:0005198; F:structural molecule activity; IEA.
      DR      InterPro: IPR003303; Filaggrin.
      DR      Pfam: PF03516; Filaggrin; 3.
      DR      PRINTS: PR00487; FILAGGRIN.
      FT      NON_TER      1
      FT      NON_TER      1
      SQ      SEQUENCE      465 AA; 50280 MW; C683744C5B134097 CRC64;

Query Match      76.1%; Score 70; DB 4; Length 465;
Best Local Similarity      78.9%; Pred. No. 0.0052;
Matches      15; Conservative      0; Mismatches      4; Indels      0; Gaps      0;

QY      1 SHOESTXGRSGRSGSGS 19
      ||||| |||||
DB      227 SHOESTXGRSGRSGSGS 245

RESULT 9
ID      075370      PRELIMINARY;      PRT;      322 AA.
AC      075370:
DT      01-NOV-1998 (TReMBLrel. 08, Created)
DT      01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Epidermal filaggrin (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99101527; PubMed=9886436;
RA      Girbal-Neuhausser E., Durieux U.J., Arnaud M., Dalbon P., Sebbag M.,
      Vincent C., Simon M., Senshu T., Masson-Bessiere C.,
      Jolivet-Reynaud C., Jolivet M., Seire G.;
RT      "The epitopes targeted by the rheumatoid arthritis-associated
      anti-filaggrin autoantibodies are posttranslationally generated on
      various sites of (pro)filaggrin by deamination of arginine residues.";
RL      U. Immunol. 162:585-594(1999).
DR      EMBL: AF043380; AAC23559.1; -.
DR      GO:0005198; F:structural molecule activity; IEA.
DR      InterPro: IPR003303; Filaggrin.
DR      Pfam: PF03516; Filaggrin; 2.
DR      PRINTS: PR00487; FILAGGRIN.
FT      NON_TER      1
FT      NON_TER      1
      SQ      SEQUENCE      322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match      75.0%; Score 69; DB 4; Length 322;
Best Local Similarity      77.8%; Pred. No. 0.0051;
Matches      14; Conservative      1; Mismatches      3; Indels      0; Gaps      0;

QY      1 SHOESTXGRSGRSGSGS 18
      ||||| |||||
DB      305 SHOESTXGRSGRSGSGS 322

RESULT 10
ID      09P312      PRELIMINARY;      PRT;      822 AA.
AC      09P312:
DT      01-OCT-2000 (TReMBLrel. 15, Created)
DT      01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT      01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE      Related to nucleolar phosphoprotein.
GN      B12P1.10.
OS      Neurospora crassa.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX      NCBI_TaxID=5141;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
      Nyakatura G., Mewes H.W., Mannhaupt G.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN      (2)
RP      SEQUENCE FROM N.A.
RA      German Neurospora genome project;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AL390091; CAB98213.1; -.
DR      PIR: T51049; T51049.
DR      GO:0003676; F:nucleic acid binding; IEA.
DR      InterPro: IPR005054; RNA_rec_mot.
DR      InterPro: IPR005120; Smg-4_UPF3.
DR      Pfam: PF00076; trm; 1.
DR      Pfam: PF03467; Smg4_UPF3; 1.
DR      SMART: SM00360; RRM; 1.
DR      PROSITE: PS50102; RRM; 1.
DR      PROSITE: PS00030; RRM_RNP_1; 1.
      SQ      SEQUENCE      822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match      59.8%; Score 55; DB 3; Length 822;
Best Local Similarity      58.8%; Pred. No. 2.3;
Matches      10; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

QY      3 QESTXGRSGRSGSGS 19
      ||||| |||||
DB      414 RESASGRTRGRGRGCT 430

RESULT 11

```


Q9V9Y3
ID Q9V9Y3 PRELIMINARY; PRT; 1284 AA.
AC Q9V9Y3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG11339-PA.
GN CG11339.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abri J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu B., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Chertys J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.S., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houtin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.D., Ketchum K.A.,
RA Jimali B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li U., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.F., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacile J.M.,
RA Palazozo J.M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheieler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.N., Venter J.C.,
RT The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Beisman C.M., Berman B.P., Carlson D.W., Ceiniker S.E.,
RA Clamp W.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Krommiller B.A., Marshall B., Milburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnick F.,
RA Whitfield B.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX FlyBase;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003777; AAF57145.2; -
DR FlyBase; FBgn003841; CG11339.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_41; 1.
DR PRINTS; PR00335; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00577; FERM_3; 1.
SQ SEQUENCE 1284 AA; 138966 MW; EDP4119838A6447 CRC64;
Query Match 54.3%; Score 50; DB 5; Length 1284;
Best Local Similarity 61.1%; Pred. No. 23;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 SHQSTXRSRSGRSG 18
DB 545 SDNESELSRSGRSG 562
RESULT 12
ID Q69375 PRELIMINARY; PRT; 820 AA.
AC Q69375;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative U25 protein.
OS Mouse cytomegalovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Bherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10366;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K181;
RX MEDLINE=94233727; PubMed=7513920;
RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;
RT "Identification and characterization of a murine cytomegalovirus gene
RT with homology to the U25 open reading frame of human
RT cytomegalovirus.";
RL Virology 200; 643-650 (1994).
DR EMBL; U02500; AAA19449.1; -
DR InterPro; IPR006731; Herpes_pp85.
DR Pfam; PF04637; Herpes_pp85; 1.
SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;
Query Match 53.3%; Score 49; DB 12; Length 820;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 QESTXGRSGRGS 14
DB 186 QSTXGRSGRGS 197
RESULT 13
ID Q9CTW2 PRELIMINARY; PRT; 849 AA.
AC Q9CTW2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F1AC21.55
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;

Tue Sep 28 07:11:24 2004

us-09-308-150-1.open.rag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92
Sequence: 1 SHQESTXGRSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_29uand4:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	19	2	AAW61515
2	90	97.8	19	2	AAW61517
3	90	97.8	19	2	AAW61505
4	90	97.8	19	2	AAW61514
5	90	97.8	19	2	AAW61516
6	86	93.5	1467	5	ABB37805
7	84	91.3	19	2	AAW61506
8	84	91.3	19	2	AAW61508
9	84	91.3	19	2	AAW61507
10	84	91.3	19	2	AAW61512
11	84	91.3	19	2	AAW61509
12	84	91.3	19	2	AAW61511
13	84	91.3	19	2	AAW61513
14	84	91.3	19	2	AAW61510
15	81	88.0	21	2	AAW61520
16	81	88.0	22	4	AAW61520
17	59	64.1	330	2	AAW61525
18	59	64.1	330	2	AAW61525
19	59	64.1	330	2	AAW61525
20	55	59.8	330	2	AAW61525
21	55	59.8	330	2	AAW61525
22	51	55.4	1771	4	AAW79819
23	50	54.3	1951	4	AAW78935
24	50	54.3	477	6	ABO07142
25	49	53.3	1199	4	ABB58274
			641	4	ABG19110

26	47	51.1	184	3	AAB57041	Aab57041 Human pro
27	46	50.0	125	2	AAY04858	Aay04858 Mycobacte
28	46	50.0	128	2	AAY04854	Aay04854 Mycobacte
29	46	50.0	443	4	AAM40069	Aam40069 Human pol
30	46	50.0	574	4	AAM41855	Aam41855 Human pol
31	46	50.0	700	4	AAE02421	Aae02421 Dario rel
32	46	50.0	752	4	AAU23535	Aau23535 Novel hum
33	45.5	49.5	135	4	ABB66181	Abb66181 Drosophi1
34	44	47.8	73	6	AAU54919	Aau54919 Propioni
35	44	47.8	128	3	ABM51438	Abm51438 Propioni
36	44	47.8	128	3	ABM51438	Abm51438 Propioni
37	44	47.8	229	7	ADBS9825	Adbs9825 Rat Prote
38	44	47.8	229	7	ADBS9833	Adbs9833 Rat Prote
39	44	47.8	229	7	ADBS9837	Adbs9837 Rat Prote
40	44	47.8	229	7	ADBS9829	Adbs9829 Rat Prote
41	44	47.8	356	5	AAE15535	Aae15535 Beta vulg
42	44	47.8	1038	7	ADC03412	Adc03412 Rice flow
43	43	46.7	92	4	AAU60843	Aau60843 Propioni
44	43	46.7	92	6	ABM57362	Abm57362 Propioni
45	43	46.7	106	3	AAU33285	Aau33285 Zea may

ALIGNMENTS

RESULT 1

AAW61515
ID AAW61515 standard; peptide; 19 AA.

AAW61515;
26-OCT-1998 (first entry)

Peptide cFA, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

solid phase synthesis; peptide amide; polyclonal antibody;

monoclonal antibody.

Synthetic.

Homo sapiens.

WO9822503-A2.

28-MAY-1998.

14-NOV-1997; 97MO-NL000624.

15-NOV-1996; 96NL-01004539.

(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

(TEME-) STICHTING TECH WETENSCHAPPEN.

Van Venrooij WJW, Scheilekens GA, Raats JMH, Hoet RMA;

WPI, 1998-398613/34.

Peptide derived from an antigen recognised by autoantibodies - is

reactive with autoimmune antibodies from rheumatoid arthritis, and may be

used in diagnosis of the disease.

Disclosure, Page 6, 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

the profilaggrin antigen which is recognised by autoantibodies from

RA patients with rheumatoid arthritis (RA). This peptide is reactive with a

RA patient's autoimmune antibodies which are reactive with profilaggrin.

The peptides were created by using standard solid phase synthesis, which

produced them as peptide amides. These sequences may be used in the

detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA;

Query Match 98.9%; Score 91; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 3.5e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 2
 AAM61517 standard; peptide; 19 AA.

AC AAM61517;
 DT 26-OCT-1998 (first entry)
 XX Peptide cfo, based on cDNA of a proflilagrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX W09822503-A2.
 PN 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

QY Sequence 19 AA;
 XX
 XX Query Match 97.8%; Score 90; DB 2; Length 19;
 XX Best Local Similarity 94.7%; Pred. No. 5e-07;
 XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 3
 AAM61505 standard; peptide; 19 AA.

AC AAM61505;

XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfo1, based on cDNA of a proflilagrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX W09822503-A2.
 PN 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

QY Sequence 19 AA;
 XX
 XX Query Match 97.8%; Score 90; DB 2; Length 19;
 XX Best Local Similarity 100.0%; Pred. No. 5e-07;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 4
 AAM61514 standard; peptide; 19 AA.

AC AAM61514;
 DT 26-OCT-1998 (first entry)
 XX Peptide cf, based on cDNA of a proflilagrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.

```

XX  WO9822503-A2.
XX
XX  28-MAY-1998.
XX
PF  14-NOV-1997; 97WO-NL000624.
XX
PR  15-NOV-1996; 96NL-01004539.
XX
PA  (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX  (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI  Van Venrooij WJM, Schellekens GA, Raats JMH, Hoeft RMA;
XX  WPI; 1998-398613/34.
XX
PT  Peptide derived from an antigen recognised by autoantibodies - is
PT  reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT  used in diagnosis of the disease.
PS  Disclosure; Page 6, 19pp; English.
XX
XX  Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC  the profilaggrin antigen which is recognised by autoantibodies from
CC  patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC  RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC  The peptides were created by using standard solid phase synthesis, which
CC  produced them as peptide amides. These sequences may be used in the
CC  detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC  for obtaining polyclonal and monoclonal antibodies
XX
SQ  Sequence 19 AA;

Query Match      97.8%; Score 90; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 5e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 SHOESTXGRSRGRSGSGS 19
    ||||| ||||| ||||| |||||
DB  1 SHOESTGRSRGRSGSGS 19

RESULT 5
AAW61516
ID  AAW61516 standard; peptide; 19 AA.
XX
AC  AAW61516;
XX
DT  26-OCT-1998 (first entry)
XX
XX  Peptide cFE, based on cDNA of a profilaggrin repeat.
XX
DE  Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX  solid phase synthesis; peptide amide; polyclonal antibody;
XX  monoclonal antibody.
XX
OS  Synthetic.
OS  Homo sapiens.
XX
PN  WO9822503-A2.
XX
PD  28-MAY-1998.
XX
PF  14-NOV-1997; 97WO-NL000624.
XX
PR  15-NOV-1996; 96NL-01004539.
XX
PA  (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX  (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI  Van Venrooij WJM, Schellekens GA, Raats JMH, Hoeft RMA;
XX  WPI; 1998-398613/34.
XX
DR

```

```

XX  Peptide derived from an antigen recognised by autoantibodies - is
XX  reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX  used in diagnosis of the disease.
XX
PS  Disclosure; Page 6, 19pp; English.
XX
XX  Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC  the profilaggrin antigen which is recognised by autoantibodies from
CC  patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC  RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC  The peptides were created by using standard solid phase synthesis, which
CC  produced them as peptide amides. These sequences may be used in the
CC  detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC  for obtaining polyclonal and monoclonal antibodies
XX
SQ  Sequence 19 AA;

Query Match      97.8%; Score 90; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 5e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 SHOESTXGRSRGRSGSGS 19
    ||||| ||||| ||||| |||||
DB  1 SHOESTGRSRGRSGSGS 19

RESULT 6
ABB97605
ID  ABB97605 standard; protein; 1467 AA.
XX
AC  ABB97605;
XX
DT  27-JUN-2002 (first entry)
XX
XX  Novel human protein SEQ ID NO: 873.
XX
DE  Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;
XX  antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
XX  neuroprotective; antiparkinsonian; protein therapy; EST;
XX  expressed sequence tag.
XX
OS  Homo sapiens.
XX
PN  WO200222660-A2.
XX
PD  21-MAR-2002.
XX
PF  10-SEP-2001; 2001WO-US026015.
XX
PR  11-SEP-2000; 2000US-00659671.
XX
PA  (HSE-) HSESEQ INC.
XX
PI  Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX  Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX  WPI; 2002-292408/33.
XX  N-PSDB; AEN32791.
XX
DR  An isolated polynucleotide for treating diseases associated with its
XX  encoded polypeptide such as cancer and multiple sclerosis.
XX
PS  Example 2; SEQ ID NO 873; 509pp; English.
XX
XX  The present invention provides the protein and coding sequences of 444
CC  novel human proteins. These were isolated from expressed sequences tags
CC  (ESTs). They can be used to stimulate cell growth, to regulate
CC  haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC  e.g. in burn treatment, to regulate the immune system e.g. to treat
CC  multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC  infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC  and cancer, to screen for drugs, to treat inflammatory conditions e.g.

```

CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
XX
SQ Sequence 1467 AA;

Query Match 93.5%; Score 86; DB 5; Length 1467;
Best Local Similarity 89.5%; Pred. No. 0.00013;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTYGRSRGRSGSGS 19
DB 773 SHOESTRGSRGRSGSGS 791

RESULT 7
AAW61506
ID AAW61506 standard; peptide; 19 AA.

AC AAW61506;

DT 26-OCT-1998 (first entry)

XX Peptide cfc2, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers
FT Modified-site 9
FT /note= "Citruiline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
PI MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
FT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTYGRSRGRSGSGS 19
|||||

DB 1 SHOESTRGSRGRSGSGS 19

RESULT 8
AAW61508
ID AAW61508 standard; peptide; 19 AA.

AC AAW61508;

DT 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers
FT Modified-site 13
FT /note= "Citruiline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
PI MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
FT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTYGRSRGRSGSGS 19
DB 1 SHOESTRGSRGRSGSGS 19

RESULT 9

AAW61507
ID AAW61507 standard; peptide; 19 AA.

AC AAW61507;

DT 26-OCT-1998 (first entry)

XX

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 11 /note= "Citruilline"
 XX
 XX MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSGRSGRSGS 19
 DB 1 SHOESTYGRSGRSGRSGS 19
 RESULT 10
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 AC AAW61512;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc8, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT
 XX
 XX MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSGRSGRSGS 19
 DB 1 SHOESTYGRSGRSGRSGS 19
 RESULT 11
 AAW61509
 ID AAW61509 standard; peptide; 19 AA.
 XX
 AC AAW61509;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 16 /note= "Citruilline"
 XX
 XX MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX MPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 SHQESTXGRSRGRSGSGS 19
DB 1 SHQESTXGRSRGRSGSGS 19
XX
RESULT 12
AAW61511
ID AAW61511 standard; peptide; 19 AA.
XX
AC AAW61511;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX MPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 4.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHQESTXGRSRGRSGSGS 19
DB 1 SHQESTXGRSRGRSGSGS 19
XX
RESULT 13
AAW61513
ID AAW61513 standard; peptide; 19 AA.
XX
AC AAW61513;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc9, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 16 /note= "Citruilline"
FT Modified-site 16 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX MPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

```


CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGRSGSGS 19
 1 SHQESTXGSRGRSGSGS 19
 Db 1 SHQESTXGSRGRSGSGS 19

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 XX
 XX 26-OCT-1998 (first entry)

DE Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

FN Key Location/Qualifiers
 FT Modified-site 7 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"

FT Modified-site 9 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"

PN WO9822503-A2.
 PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA,
 DR WPI, 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGRSGSGS 19
 1 SHQESTXGSRGRSGSGS 19
 Db 1 SHQESTXGSRGRSGSGS 19

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 XX
 XX 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

FN Key Location/Qualifiers
 FT Modified-site 9 /note= "Citnulline"

FT Modified-site 9 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"

PN WO9822503-A2.
 PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA,
 DR WPI, 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 21 AA;

Query Match 88.0%; Score 81; DB 2; Length 21;
 Best Local Similarity 94.4%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HQESTXGSRGRSGSGS 19
 4 HQESTXGSRGRSGSGS 21
 Db 4 HQESTXGSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:09
 Job time : 57.1302 secs

Blank Sheet

Tue Sep 28 07:33:33 2004

us-09-308-150-1.open.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHOESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321556718 residues 1349238

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	19	9	US-09-308-150-12 Sequence 12, Appl
2	90	97.8	19	9	US-09-308-150-11 Sequence 11, Appl
3	90	97.8	19	9	US-09-308-150-11 Sequence 11, Appl
4	90	97.8	19	9	US-09-308-150-13 Sequence 13, Appl
5	90	97.8	19	9	US-09-308-150-14 Sequence 14, Appl
6	84	91.3	19	9	US-09-308-150-2 Sequence 2, Appl
7	84	91.3	19	9	US-09-308-150-3 Sequence 3, Appl
8	84	91.3	19	9	US-09-308-150-4 Sequence 4, Appl
9	84	91.3	19	9	US-09-308-150-5 Sequence 5, Appl
10	84	91.3	19	9	US-09-308-150-6 Sequence 6, Appl
11	84	91.3	19	9	US-09-308-150-7 Sequence 7, Appl
12	84	91.3	19	9	US-09-308-150-8 Sequence 8, Appl
13	84	91.3	19	9	US-09-308-150-9 Sequence 9, Appl
14	81	88.0	21	9	US-09-308-150-10 Sequence 10, Appl
15	81	88.0	22	9	US-09-747-0239A-22 Sequence 22, Appl

16	50	54.3	477	15	US-10-161-927-62	Sequence 62, Appl
17	50	54.3	1087	9	US-09-918-909-24	Sequence 24, Appl
18	50	54.3	1087	16	US-10-641-991-24	Sequence 24, Appl
19	48	52.2	179	12	US-10-425-114-53545	Sequence 53545, A
20	48	52.2	436	15	US-10-369-493-8178	Sequence 8178, Ap
21	48	52.2	506	16	US-10-437-963-185974	Sequence 185974, A
22	47	51.1	50	14	US-10-029-386-27692	Sequence 27692, A
23	47	51.1	123	16	US-10-767-701-32436	Sequence 32436, A
24	47	51.1	124	9	US-09-925-300-1619	Sequence 1619, Ap
25	46	50.0	164	16	US-10-437-963-112419	Sequence 112419, A
26	46	50.0	384	12	US-10-425-114-72136	Sequence 72136, A
27	46	50.0	384	12	US-10-425-114-72137	Sequence 72137, A
28	46	50.0	436	14	US-10-156-761-13022	Sequence 13022, A
29	46	50.0	580	12	US-10-425-114-70952	Sequence 70952, A
30	46	50.0	612	12	US-10-424-599-230819	Sequence 230819, A
31	45	48.9	223	16	US-10-437-963-133033	Sequence 133033, A
32	45	48.9	838	16	US-10-437-963-146503	Sequence 146503, A
33	45	48.9	854	12	US-10-425-114-57838	Sequence 57838, A
34	44.5	48.4	85	12	US-10-424-559-161928	Sequence 161928, A
35	44	47.8	52	16	US-10-437-963-155141	Sequence 155141, A
36	44	47.8	76	12	US-10-424-599-248912	Sequence 248912, A
37	44	47.8	78	16	US-10-437-963-203073	Sequence 203073, A
38	44	47.8	105	16	US-10-437-963-196208	Sequence 196208, A
39	44	47.8	131	16	US-10-767-701-33185	Sequence 33185, A
40	44	47.8	232	16	US-10-437-963-183394	Sequence 183394, A
41	44	47.8	262	12	US-10-425-114-72166	Sequence 72166, A
42	44	47.8	262	12	US-10-425-114-72167	Sequence 72167, A
43	44	47.8	340	12	US-10-425-114-68797	Sequence 68797, A
44	44	47.8	406	12	US-10-425-114-41013	Sequence 41013, A
45	44	47.8	843	14	US-10-359-431-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1

US-09-308-150-12

Sequence 12, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

US-09-308-150-12

Query Match 98.9%; Score 91; DB 9; Length 19;
Best Match Similarity 94.7%; Pred No. 2.9e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19

Db 1 SHOESTXGRSRGRSGS 19

RESULT 2

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; OTHER INFORMATION: Xaa is citrulline
 ; US-09-308-150-1

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGS 19
 Db 1 SHOESTXGRSRGRSGS 19

RESULT 3

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19

TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4e-06;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGS 19
 Db 1 SHOESTXGRSRGRSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; US-09-308-150-13

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGS 19
 Db 1 SHOESTXGRSRGRSGS 19

RESULT 5

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-09-308-150-14
OTHER INFORMATION: known cDNA sequences of human profilaggrin
```

```
Query Match      97.8%  Score 90; DB 9; Length 19;
Best Local Similarity 94.7%  Pred. No. 4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 6
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-2
```

```
Query Match      91.3%  Score 84; DB 9; Length 19;
Best Local Similarity 89.5%  Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 7
US-09-308-150-3
Sequence 3, Application US/09308150
```

```
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-3
```

```
Query Match      91.3%  Score 84; DB 9; Length 19;
Best Local Similarity 89.5%  Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 8
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-4
```

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 9
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 10
US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 11
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 12
US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8
```

```
Query Match      91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTXGRSRGRSGSGS 19
```

```
RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9
```

```
Query Match      91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTXGRSRGRSGSGS 19
```

```
RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
NAME/KEY: DISULFID
LOCATION: (3)..(16)
US-09-308-150-10
```

```
Query Match      88.0%; Score 81; DB 9; Length 21;
Best Local Similarity 94.4%; Pred. No. 8.3e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      2 HOESTXGRSRGRSGSGS 19
Db      4 HOESTXGRSRGRSGSGS 21
```

```
RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydie
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.003INP500 INNS:031
CURRENT FILING DATE: 2000-12-21
CURRENT APPLICATION NUMBER: US/09/747,029A
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-1
Perfect score: 92
Sequence: 1 SHQESTXGSRGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	51.1	255	4	US-09-252-991A-19598
2	47	51.1	421	4	US-09-252-991A-32326
3	45	48.9	491	4	US-09-489-039A-7836
4	44	47.8	518	4	US-09-252-991A-25967
5	44	47.8	843	4	US-08-591-502B-45
6	43.5	47.3	326	4	US-09-252-991A-17002
7	43	46.7	239	4	US-09-252-991A-28985
8	43	46.7	395	4	US-09-252-991A-21702
9	43	46.7	483	4	US-09-252-991A-19015
10	43	46.7	546	4	US-09-252-991A-24620
11	43	46.7	653	4	US-09-513-057C-13
12	43	46.7	722	4	US-09-884-890-4
13	43	46.7	724	4	US-09-884-890-2
14	43	46.7	730	4	US-08-591-502B-49
15	43	46.7	745	4	US-09-523-849-36
16	43	46.7	842	4	US-08-591-502B-50
17	43	46.7	842	4	US-08-591-502B-51
18	43	46.7	843	4	US-09-719-528A-2
19	43	46.7	843	4	US-08-591-502B-48
20	43	46.7	843	4	US-08-591-502B-59
21	43	46.7	845	4	US-08-591-502B-46
22	42	45.7	138	4	US-09-252-991A-16961
23	42	45.7	169	3	US-09-342-084-6
24	42	45.7	409	4	US-09-489-039A-14217
25	42	45.7	408	4	US-09-252-991A-24066
26	42	45.7	629	4	US-09-252-991A-22901
27	42	45.7	708	4	US-09-252-991A-18074

28	42	45.7	722	4	US-08-817-832B-32	Sequence 32, Appl
29	41	44.6	85	4	US-09-252-991A-29339	Sequence 29339, A
30	41	44.6	180	4	US-09-252-991A-23744	Sequence 23744, A
31	41	44.6	204	4	US-09-252-991A-21317	Sequence 21317, A
32	41	44.6	236	4	US-09-252-991A-25980	Sequence 25980, A
33	41	44.6	256	4	US-09-252-991A-26244	Sequence 26244, A
34	41	44.6	258	4	US-09-252-991A-28812	Sequence 28812, A
35	41	44.6	294	4	US-09-252-991A-21947	Sequence 21947, A
36	41	44.6	363	4	US-09-252-991A-25132	Sequence 25132, A
37	41	44.6	551	4	US-09-252-991A-28829	Sequence 28829, A
38	41	44.6	586	4	US-09-252-991A-24514	Sequence 24514, A
39	41	44.6	2237	1	US-08-354-973-1	Sequence 1, Appl
40	40.5	44.0	96	4	US-09-252-991A-19216	Sequence 19216, A
41	40.5	44.0	415	4	US-08-328-352-6430	Sequence 6430, Ap
42	40	43.5	120	4	US-09-252-991A-17025	Sequence 17025, A
43	40	43.5	149	4	US-09-252-991A-24582	Sequence 24582, A
44	40	43.5	183	4	US-09-252-991A-25616	Sequence 25616, A
45	40	43.5	195	4	US-09-252-991A-30009	Sequence 30009, A

ALIGNMENTS

```

RESULT 1
US-09-252-991A-19598
; Sequence 19598, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19598
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19598

Query Match      51.1% Score 47; DB 4; Length 255;
Best Local Similarity 66.7%; Pred. No. 5.9;
Matches      8; Conservative      3; Mismatches      1; Indels      0; Gaps      0;

Cy      8 GSRGRSGRSGS 19
Db      64 GRDRGRAGRAGA 75

RESULT 2
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT

```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25967

Query Match 51.1%; Score 47; DB 4; Length 421;
Best Local Similarity 52.9%; Pred. No. 9.8;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGSG 18
DB 13 HARTGTGTGRCGRSG 29

RESULT 3
US-09-489-039A-7836
Sequence 7836, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7836
LENGTH: 491

TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7836

Query Match 48.9%; Score 45; DB 4; Length 491;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 STXGRSGRSGSGS 17
DB 217 TAGSRSGRSGSGS 229

RESULT 4
US-09-252-991A-25967
Sequence 25967, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 25967
LENGTH: 518
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25967

Query Match 47.8%; Score 44; DB 4; Length 518;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 STXGRSGRSGSG 18
DB 412 SGGSGRSGRSGSG 425

RESULT 5

US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus

NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898

FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
APPLICATION NUMBER: WO PCT/US94/08685

FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver

REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45

Query Match 47.8%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 57;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGSGS 19
DB 226 QPQGSMAKSGRSGS 242

RESULT 6
US-09-252-991A-17002
Sequence 17002, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 17002
;; LENGTH: 326
;; TYPE: PR1
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17002

Query Match 47.3%; Score 43.5; DB 4; Length 326;
Best Local Similarity 68.8%; Pred. No. 26;
Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 4 ESTXGRSGRSGSGS 19
Db 236 ESSGSRARS-RDGS 250

RESULT 7
US-09-252-991A-28985
; Sequence 28985, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28985
; LENGTH: 239
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28985

Query Match 46.7%; Score 43; DB 4; Length 239;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGRS 17
Db 171 SHDRAPGRSRDRAGRA 167

RESULT 8
US-09-252-991A-21702
; Sequence 21702, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21702
; LENGTH: 395
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21702

Query Match 46.7%; Score 43; DB 4; Length 395;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGRS 18
Db 251 SHQASGRGDPHESGRPG 268

RESULT 9
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015

Query Match 46.7%; Score 43; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQESTXGRSGRSGRS 18
Db 94 HAEODPGAGTGRAGRAG 110

RESULT 10
US-09-252-991A-24620
; Sequence 24620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24620
; LENGTH: 546
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24620

Query Match 46.7%; Score 43; DB 4; Length 546;
Best Local Similarity 64.3%; Pred. No. 52;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGRSG 18
Db 427 SRLGRDRGASGRPG 440

RESULT 11
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:

APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 653
TYPE: PRT
ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 46.7%; Score 43; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 62;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SHORSTGRSGRSGSGS 19
Db 573 SRQVSTASASAGRGISGS 591

RESULT 12
US-09-984-890-4
Sequence 4, Application US/09984890
Patent No. 6492156
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C0001306
CURRENT APPLICATION NUMBER: US/09/984,890
CURRENT FILING DATE: 2001-10-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 722
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match 46.7%; Score 43; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 5 STGRSGRSGRSGS 19
Db 562 SPSGHSQGRSGS 576

RESULT 13
US-09-984-890-2
Sequence 2, Application US/09984890
Patent No. 6492156
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C0001306
CURRENT APPLICATION NUMBER: US/09/984,890
CURRENT FILING DATE: 2001-10-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 724
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-890-2

Query Match 46.7%; Score 43; DB 4; Length 724;
Best Local Similarity 60.0%; Pred. No. 69;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 5 STGRSGRSGRSGS 19
Db 564 SPSGHSQGRSGS 578

RESULT 14
US-08-591-502B-49
Sequence 49, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytoprotective T
Lymphocyte Responses to Hepatitis B Virus
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 99
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-591-502B-49

Query Match 46.7%; Score 43; DB 4; Length 730;
Best Local Similarity 80.0%; Pred. No. 70;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 10 SRGRSGRSGS 19
Db 233 ARKSGRSGS 242

RESULT 15
US-09-523-849-36
Sequence 36, Application US/09523849
Patent No. 6458551
GENERAL INFORMATION:
APPLICANT: Bandman, Olga

```

; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hoegson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523, 849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36

```

```

Query Match          46.7%; Score 43; DB 4; Length 745;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 5 STXGRSRGRGRSGS 19
Db 585 SPSGHSQGRGRGASGS 599

```

Search completed: September 28, 2004, 06:26:40
 Job time : 15.4479 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHQESTRGXSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	97.8	2248	2	A35938	profilaggrin - hum
2	84	91.3	416	2	A32947	filaggrin precursor
3	75	81.5	591	2	A45135	profilaggrin - hum
4	58	63.0	506	1	W2WL47	E2 protein - human
5	50	54.3	800	2	T02852	probable membrane
6	48	52.2	822	2	T51049	related to nucleol
7	47	51.1	471	2	T33997	hypothetical prote
8	46	50.0	197	2	T18918	hypothetical prote
9	46	50.0	203	2	C87801	protein C10G1.9 l
10	46	50.0	399	2	T46259	hypothetical prote
11	45	48.9	836	2	G84727	probable DNA topoi
12	45	48.9	374	2	T33328	hypothetical prote
13	45	48.9	1337	2	T30291	hypothetical prote
14	45	48.9	1791	2	T02345	dextranase - Strept
15	44	47.8	229	2	UC7219	hypothetical prote
16	44	47.8	790	2	T05576	nuclear protein SR
17	44	47.8	1232	2	S40766	hypothetical prote
18	43.5	47.3	463	2	T51194	hypothetical prote
19	43	46.7	123	2	T16234	hypothetical prote
20	43	46.7	409	2	T35118	hypothetical prote
21	43	46.7	452	2	S35482	probable secreted
22	43	46.7	525	2	T35194	E2 protein - human
23	43	46.7	827	2	T13468	cdc35A - rat
24	43	46.7	842	1	JDVLVS	DNA-directed DNA p
25	43	46.7	843	1	JDVLVR	DNA-directed DNA p
26	42	45.7	203	2	T25916	hypothetical prote
27	42	45.7	306	2	T21220	hypothetical prote
28	42	45.7	694	2	S71785	wingless receptor
29	42	45.7	849	2	A96592	hypothetical prote

30	42	45.7	889	2	T33422	hypothetical prote
31	42	45.7	1287	2	S55954	viral mRNA transla
32	42	45.7	1829	2	T35681	probable sensory h
33	42	45.7	2233	2	S63347	acetyl-CoA carboxy
34	41.5	45.1	646	2	H96665	protein F2C12.10
35	41.5	45.1	1895	2	T06609	disease resistance
36	41	44.6	136	2	T35632	probable transposa
37	41	44.6	200	1	OCB827	ERLPI protein - hu
38	41	44.6	205	1	S19634	nucleolar protein
39	41	44.6	312	2	A31846	130K paracrystalli
40	41	44.6	315	2	S53589	SOL2 protein - yea
41	41	44.6	360	2	T14444	hypothetical prote
42	41	44.6	373	2	T02976	probable DNA bindi
43	41	44.6	408	2	T21412	hypothetical prote
44	41	44.6	458	2	G83690	hypothetical prote
45	41	44.6	461	2	S36593	E2 protein - human

ALIGNMENTS

```
RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GAM>
A:Cross-references: GB:J02929
A:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
P:246-569/Region: filaggrin repeat
F:570-893/Region: filaggrin repeat
F:1074-1397/Region: filaggrin repeat
F:1573-1896/Region: filaggrin repeat

Query Match          97.8% Score 90; DB 2; Length 2248;
Best Local Similarity 94.7%; Pred. No. 8.9e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 SHQESTRGXSRGSRGSGS 19
DB      551 SHQESTRGXSRGSRGSGS 569

RESULT 2
A32947
filaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cammizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A:Title: Characterization of a cDNA clone encoding human filaggrin and localization of tr
A:Reference number: A32947; MUID:99296901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M2435; NID:g182604; PIND:AA52454.1; PID:g182605
A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
```

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 91.3%; Score 84; DB 2; Length 416;
 Best Local Similarity 89.5%; Pred. No. 1.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 |||||
 7 SHQSTRGXSRGSRGSGS 25

RESULT 3
 A45135
 A:Protein: profilaggrin - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135
 R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nixonsukisiri, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992
 A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717
 A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101089; GB:M90967; NID:G190408; PIDN:AAA60177.1; PID:9553621
 A:Note: sequence extracted from NCBI backbone (NCBI:FP.118773)
 C:Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 81.5%; Score 75; DB 2; Length 591;
 Best Local Similarity 84.2%; Pred. No. 0.00051;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 |||||
 449 SHQSTRGXSRGSRGSGS 467

RESULT 4
 M2ML47
 A:Protein: human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D5324
 R:Kiyono, T.; Adachi, A.; Ishibashi, M.
 Virology 177, 401-405, 1990
 A:Title: Genome organization and taxonomic position of human papillomavirus type 47
 A:Reference number: A5324; MUID:90281611; PMID:2162112
 A:Accession: D5324
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KTY>
 A:Cross-references: GB:M2305; NID:G333062; PIDN:AAA46979.1; PID:G333067
 C:Superfamily: papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.19;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 | : : ||| ||| ||| : ||

DB 342 SRENTRGXSRGSRGSGS 360

RESULT 5
 T02852
 A:Protein: Probable membrane protein U1439.4 [imported] - Leishmania major (strain Friedlin)
 C:Species: Leishmania major
 C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C:Accession: T02852; H81462
 R:Myler, P.J.
 submitted to the EMBL Data Library, May 1998
 A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
 A:Reference number: 214740
 A:Accession: T02852
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-800 <MYL>
 A:Cross-references: EMBL:AE001274; NID:G3264850; PID:G2266920
 R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A:Reference number: A81455; MUID:99178987; PMID:10077609
 A:Accession: H81462
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-800 <PYL>
 A:Cross-references: GB:AE001274; NID:G3264850; PIDN:AA024675.1; PID:G2266920; GSPDB:GN001
 A:Experimental source: strain MEOM/IL/81/Friedlin
 C:Genetics:
 A:Gene: U1439.4
 A:Map position: 1

Query Match 54.3%; Score 50; DB 2; Length 800;
 Best Local Similarity 55.6%; Pred. No. 5.2;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 HQSTRGXSRGSRGSGS 19
 ||: ||| |||||
 429 HRDGVGRGSRGSRGSGS 446

RESULT 6
 T51049
 A:Protein: related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12FL10
 C:Species: Neurospora crassa
 C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049
 R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: 225286
 A:Accession: T51049
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12FL1.10
 A:Experimental source: BAC clone B12FL1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B12FL1.10
 A:Map position: 6
 A:introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 11;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTRGXSRGSRGSGS 19
 : ||| : ||| ||| : |||
 DB 414 RBSAGRTGRGRGSGT 430

RESULT 7
 T33997

hypothetical protein W03G1.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T33997
 R/Pauley, A.; Schreel, P.; Harper, M.
 submitted to the EMBL Data Library, February 1999
 A/Description: The sequence of C. elegans cosmid W03G1.
 A/Reference number: Z21454
 A/Accession: T33997
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-471 <PAU>
 A/Cross-references: EMBL:AF125964; PIDN:AD14753.1; GSPDB:GN00022; CESP:W03G1.5
 A/Experimental source: strain Bristol N2; clone W03G1
 C/Genetics:
 A/Gene: CESP:W03G1.5
 A/Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
 Best Local Similarity 52.6%; Pred. No. 9;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSGSGS 19
 DB 167 SRSRSGSVGRGSGSGS 185

RESULT 8

hypothetical protein C04G2.8 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T18918
 R/Hemby, C.
 submitted to the EMBL Data Library, April 1996
 A/Reference number: Z19045
 A/Accession: T18918
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-197 <WIL>
 A/Cross-references: EMBL:Z70718; PIDN:CAA94670.1; GSPDB:GN00022; CESP:C04G2.8
 A/Experimental source: clone C04G2
 C/Genetics:
 A/Gene: CESP:C04G2.8
 A/Map position: 4
 A/Introns: 70/2

Query Match 50.0%; Score 46; DB 2; Length 197;
 Best Local Similarity 52.6%; Pred. No. 5.5;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSGSGS 19
 DB 104 SRSRSGSVGRGSGSGS 122

RESULT 9

protein C10G11.9 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C/Accession: C87801
 R/anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A/Reference number: A75000; MUID:99069613; PMID:9851916
 A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A/Note: Published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A/Accession: C87801
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-203 <STO>
 A/Cross-references: GB:chr_I, PIDN:AA42250.1; PID:q1825626; GSPDB:GN00019; CESP:C10G11.

C/Genetics:
 A/Gene: C10G11.9
 A/Map position: 1

Query Match 50.0%; Score 46; DB 2; Length 203;
 Best Local Similarity 52.6%; Pred. No. 5.6;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSGSGS 19
 DB 110 SRSRSGSVGRGSGSGS 128

RESULT 10

hypothetical protein DKFZp761E0323.1 - human (fragment)

C/Species: Homo sapiens (man)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
 C/Accession: T46259
 R/Bloecher, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A/Reference number: Z23032
 A/Accession: T46259
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-399 <AAA>
 A/Cross-references: EMBL:AL137423
 A/Experimental source: adult amygdala; clone DKFZp761E0323
 C/Genetics:
 A/Note: DKFZp761E0323.1

Query Match 50.0%; Score 46; DB 2; Length 399;
 Best Local Similarity 55.6%; Pred. No. 11;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSGSGS 18
 DB 336 SRSRSGSVGRGSGSGS 353

RESULT 11

probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: G84727
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: G84727
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-836 <STO>
 A/Cross-references: GB:AE002093; NID:94263718; PIDN:AD15404.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2g32000
 A/Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
 Best Local Similarity 69.2%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGXSRGSGSGSGS 19
 DB 812 RGRGRGGRGSGSGS 824

RESULT 12

hypothetical protein F13C5.2 - Caenorhabditis elegans

C/Species: *Caenorhabditis elegans*
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Sep-2000
 C/Accession: T33328
 R/Nohlmann, P.; Hawkins, J.; Gillam, B.
 Submitted to the EMBL Data Library, July 1998
 A/Description: The sequence of *C. elegans* cosmid F13C5.
 A/Reference number: Z21324
 A/Accession: T33328
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-374 <NOH>
 A/Cross-references: EMBL:AF077531; PIDN:AA04610.1; GSPDB:GN00028; CESP:F13C5.2
 A/Experimental source: strain Bristol N2; clone F13C5
 C/Genetics:
 A/Gene: CESP:F13C5.2
 A/Map position: X
 A/Intons: 25/3; 135/3; 189/3; 313/1
 C/Superfamily: bromodomain homology
 F141-198/Domain: bromodomain homology <BRO>

Query Match 48.9%; Score 45; DB 2; Length 374;
 Best Local Similarity 69.2%; Pred. No. 15;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ESTRGXSRGSGR 16
 DB 64 ESTRGCTGRGGR 76

RESULT 13
 T30291
 dextransase - *Streptococcus sobrinus*
 C/Species: *Streptococcus sobrinus*
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30291
 R/Wanda, S.Y.; Curtiss III, R.
 J. Bacteriol. 176, 3839-3850, 1994
 A/Title: Purification and characterization of *Streptococcus sobrinus* dextransase produced
 A/Reference number: Z20810; MUID:94292401; PMID:80211165
 A/Accession: T30291
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1337 <MAN>
 A/Cross-references: EMBL:M96978; NID:9450640; PID:9450641; PIDN:AAA21772.1

Query Match 48.9%; Score 45; DB 2; Length 1337;
 Best Local Similarity 47.1%; Pred. No. 51;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTRGXSRGSGRS 17
 DB 1154 ANQSTKGSADQSGKS 1170

RESULT 14
 T02345
 hypothetical protein KIAA0324 - human (fragment)
 C/Species: *Homo sapiens* (man)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
 C/Accession: T02345
 R/Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, O.; White, S.; Deng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 Submitted to the EMBL Data Library, March 1998
 A/Description: Sequencing of human chromosome 16p13.3.
 A/Reference number: Z14664
 A/Accession: T02345
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1791 <RIC>
 A/Cross-references: EMBL:AC004493; NID:92996648; PIDN:AA008453.1; PID:92996650
 C/Genetics:
 A/Map position: 16
 A/Intons: 1610/2; 1706/2

A/Note: KIAA0324

Query Match 48.9%; Score 45; DB 2; Length 1791;
 Best Local Similarity 58.8%; Pred. No. 68;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTRGXSRGSGRS 17
 DB 858 ARQSSRTSSRRRGRS 874

RESULT 15
 J07219
 nuclear protein SR-25 - mouse
 C/Species: *Mus musculus* (house mouse)
 C/Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
 C/Accession: J07219
 R/Saahara, K.; Yamaoka, T.; Moritani, M.; Tanaka, M.; Iwahana, H.; Yoshimoto, K.; Miyagi
 Biochem. Biophys. Res. Commun. 269, 444-450, 2000
 A/Title: Molecular cloning and expression analysis of a putative nuclear protein, SR-25.
 A/Reference number: J07219; MUID:20175222; PMID:10708573
 A/Accession: J07219
 A/Molecule type: mRNA
 A/Residues: 1-229 <SAS>
 A/Cross-references: DBJ:AB035383; NID:97619895; PIDN:BA094743.1; PID:97619896
 A/Experimental source: MIN6 cell line
 A/Comment: This protein is a highly hydrophilic nuclear protein with a serine-arginine re
 A splicing factors
 A/Keywords: nucleus; RNA processing

Query Match 47.8%; Score 44; DB 2; Length 229;
 Best Local Similarity 52.9%; Pred. No. 13;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 CESTRGXSRGSGRS 19
 DB 7 RKRSRSRSGSGRS 23

Search completed: September 28, 2004, 06:14:57
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHQESTRGSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	91.3	416	1	FILA_HUMAN
2	58	63.0	506	1	VE2_HPV47
3	45	48.9	1337	1	DEXT_STRDO
4	44	47.8	1232	1	YQO3
5	43	46.7	123	1	LSM4_CAEEL
6	43	46.7	452	1	VE2_HPV17
7	43	46.7	525	1	MRP1_RAT
8	43	46.7	730	1	DPO1_HPBV4
9	43	46.7	842	1	DPO1_HPBV7
10	43	46.7	843	1	DPO1_HPBV8
11	42.5	46.2	388	1	ROG_MOUSE
12	42.5	46.2	391	1	ROG_HUMAN
13	42.5	46.2	1180	1	TYK2_MOUSE
14	42	45.7	694	1	FRZ2_DROME
15	42	45.7	1097	1	RNT2_FUGRU
16	42	45.7	1287	1	SKT2_YEAST
17	42	45.7	2233	1	COAC_YEAST
18	41.5	45.1	1895	1	WR19_ARATH
19	41	44.6	197	1	RX21_DROME
20	41	44.6	205	1	GAR1_YEAST
21	41	44.6	315	1	SOL2_YEAST
22	41	44.6	461	1	VE2_HPV09
23	41	44.6	503	1	VE2_HPV21
24	41	44.6	772	1	WR11_CAEEL
25	41	44.6	774	1	MRK2_MOUSE
26	41	44.6	825	1	SE5_RAT
27	41	44.6	955	1	CLIS2_HUMAN
28	41	44.6	2404	1	SON_MOUSE
29	41	44.6	2426	1	SON_HUMAN
30	40.5	44.0	255	1	THO4_MOUSE
31	40.5	44.0	257	1	THO4_HUMAN
32	40.5	44.0	1509	1	MYSN_ACACA
33	40	43.5	414	1	NSR1_YEAST

ALIGNMENTS

ID	FILE_HUMAN	STANDARD	PRT	416 AA.
AC	P20930			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Filaggrin precursor (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89296901; PubMed=2740331;			
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,			
RA	Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.,			
RT	"Characterization of a cDNA clone encoding human filaggrin and			
RT	localization of the gene to chromosome region 1q21."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).			
RN	[2]			
RP	CITRULLINATION.			
RX	MEDLINE=96374388; PubMed=8780679;			
RA	Senshu T., Kan S., Ogawa H., Manabe M., Asaga H.,			
RT	"Preferential delamination of keratin K1 and filaggrin during the			
RT	terminal differentiation of human epidermis."			
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).			
CC	-1- FUNCTION: Aggregates keratin intermediate filaments and promotes			
CC	disulfide-bond formation among the intermediate filaments during			
CC	terminal differentiation of mammalian epidermis.			
CC	-1- PTM: Filaggrin is initially synthesized as a large, insoluble,			
CC	highly phosphorylated precursor containing many tandem copies of			
CC	324 AA, which are not separated by "large linker". The precursor			
CC	is deposited as keratohyalin granules. During terminal			
CC	differentiation it is dephosphorylated and proteolytically			
CC	cleaved.			
CC	-1- PTM: Undergoes delamination of some arginine residues			
CC	(citrullination).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M24355; AAA52454.1; -			P54672 dictyostell
DR	PIR; A32947; A32947.			P36782 human papil
DR	Gene; HGNC:3748; FLG.			P81361 mus musculu
DR	MIM; 135940; -			P22807 dtrosophila
DR	GO; GO:0005882; C:intermediate filament; NAS.			P00192 escherichia
DR	GO; GO:0005198; F:structural molecule activity; NAS.			000586 crepusculi
DR	GO; GO:0007275; P:development; NAS.			P02463 mus musculu
DR	InterPro; IPR003303; Filaggrin.			P16629 homo sapien

```

DR Pfam: PF03516; Filaggrin. 2.
DR PRINTS: PRO0487; FILAGRIN.
DR phosphorylation; Citrullination; Developmental protein.
KM NON_TER 1
SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;
OY 1 SH0ESTRGXSRGSRGSGS 19
Db 7 SH0ESTRGXSRGSRGSGS 25
Query Match 91.3%; Score 84; DB 1; Length 416;
Best Local Similarity 89.5%; Pred. No. 0;076;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
RESULT 2
ID VE2_HPv47 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE 15-Regulatory protein E2.
GN E2.
OS Human papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2B RESPONSE ELEMENT (5'-ACGNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2B'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M32305; AAA46979.1; -.
DR PIR: D35324; W2NL47.
DR HSSP: P03122; 2BOP.
DR InterPro: IPR000427; E2_C.
DR InterPro: IPR001866; E2_N.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF00511; E2_C; 1.
DR Pfam: PF00508; E2_N; 1.
DR ProDom: PD000672; E2_C; 1.
DR ProDom: PD000678; E2_N; 1.
DR Early protein; Transcription regulation; Activator; DNA-binding;
DR Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C374BF725065 CRC64;
Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.076;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
1 SH0ESTRGXSRGSRGSGS 19

```

```

Db          : 342 SRENTNRGRGRGRGRAGS 360

RESULT 3
DXT_STRDO  STANDARD;      PRT; 1337 AA.
AC          P39653;
DT          01-FEB-1995 (Rel. 31, Created)
DT          01-FEB-1995 (Rel. 31, Last sequence update)
DT          10-OCT-2003 (Rel. 42, Last annotation update)
DE          Dextranase precursor (EC 3.2.1.11) (Alpha-L-6-glucan-6-
            glucanohydrolase).
GN          DEX.
OS          Streptococcus downei (Streptococcus sobrinus).
CS          Plasmid pYA902.
OC          Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC          Streptococcus.
OK          NCBI_TaxID=1317;
            [1]
            SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
RN          RP
RP          STRAIN=6715 / UAB66;
RC          MEDLINE=94292401; PubMed=8021165;
RX          Wanda S.-Y., Curtiss R. III;
RA          "Purification and characterization of Streptococcus sobrinus
            dextranase produced in recombinant Escherichia coli and sequence
            analysis of the dextranase gene."
RL          J. Bacteriol. 176:3839-3850 (1994).
CC          -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
            PELICLE-COATED TOOTH SURFACE.
CC          -1- CATALYTIC ACTIVITY: Endonucleolysis of 1,6-alpha-D-glucosidic
            linkages in dextran.
CC          -1- SUBUNIT: Homodimer.
CC          -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
            an amide bond (potential).
CC          -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
            and at 39 degrees Celsius.
CC          -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
CC          -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
-----
CC          CC
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          CC the European Bioinformatics Institute. There are no restrictions on its
CC          CC use by non-profit institutions as long as its content is in no way
CC          CC modified and this statement is not removed. Usage by and for commercial
CC          CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
            or send an email to license@isb-sib.ch).
CC          CC
CC          -----
DR          EMBL; M96978; AAA21772.1; -.
DR          InterPro; IPR001893; Gram_pos_anchor.
DR          Pfam; PF00746; Gram_pos_anchor.1.
DR          TIGRfam; TIGR01167; LPXG_anchor.1.
DR          PROSITE; PS50847; GRAM_POS_ANCHORING.1.
KW          Hydroxylase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
            Signal; Plasmid.
FT          SIGNAL 1
FT          CHAIN 31 1308 DEXTRANASE.
FT          PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
FT          STR 1305 1309 LPXG SORTING SIGNAL (POTENTIAL).
FT          MOD RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ          SEQUENCE 1337 AA; 143298 MW; B494275A77A2B3D0 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 24;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY          1 SHQSTRGXSRGRSGRS 17
            :|:|:|:|:|:|:|:|:|
Db          1154 ANQDSTGSSADQSGKS 1170

```

```

ID Y005 CAEBL STANDARD; PRT; 1232 AA.
AC P34633;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK512.5 in chromosome III.
GN ZK512.5.
OS Caenorhabditis elegans.
OC Rhabditidae; Peleceroidea; Caenorhabditis.
OC Rhabditidae; Peleceroidea; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Atencough R., Anderson K., Baynes C., Berke M., Bonfield U., Burton J., Connell M., Copesey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R., Sims M., Smaiden N., Smith A., Smith M., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
RT RT
RL Nature 368:32-38(1994).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

CC -!- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (by similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```
DR PIR; S36482; S36482.
DR HSSP; P1783; IDHM.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR005021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C; 1.
DR Pfam; PF00508; E2_N; 1.
DR ProDom; PD000672; E2_C; 1.
DR ProDom; PD000673; E2_N; 1.
DR Early protein: Transcription regulation: Activator; DNA-binding;
Trans-acting factor; DNA replication; Repressor; Nuclear protein.
KW SEQUENCE 452 AA; 51241 MW; 70296AF16D3E55 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 452;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 QESTRGSRGSRGSG 18
DB 293 RDSRSPNRGRGSGSG 308

RESULT 7
MPI RAT STANDARD; PRT; 525 AA.
AC P46965;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE M-phase inducer phosphatase 1 (EC 3.1.3.48) (Dual specificity
phosphatase Cdc55A).
GN CDC25A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRK49F;
RX MEDLINE=94208523; PubMed=8156993;
RA Ujino S., Suto K., Nagata A., Igatahshi M., Kanaoka Y., Nojima H.,
RA Okayama H.;
RT "Cdc25A is a novel phosphatase functioning early in the cell cycle.";
RL EMBO J. 13:1549-1556(1994).
CC -1- FUNCTION: Functions as a dosage-dependent inducer in mitotic
control. It is a tyrosine protein phosphatase required for
progression of the cell cycle. It directly dephosphorylates CDC2
and activate its kinase activity.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
tyrosine + phosphate.
CC -1- SIMILARITY: Belongs to the MPI phosphatase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16236; EAA03761.1; -.
DR HSSP; P30304; 1C25.
DR InterPro; IPR000751; MPI_Phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR00716; MPIPHMTASE.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR Cell division; Mitosis; Hydrolase; Multigene family.
KW DOMAIN 377 483
FT ACT SITE 432 432 BY SIMILARITY.
SEQUENCE 525 AA; 59217 MW; A60EC0AF46D63A9C CRC64;
```

```
Query Match 46.7%; Score 43; DB 1; Length 525;
Best Local Similarity 47.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSG 17
DB 283 SHEESPRTGTRKRKSEA 299

RESULT 8
DPOI_HPBV4 STANDARD; PRT; 730 AA.
ID DPOI_HPBV4
AC P12933;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Hepatitis B virus (subtype adr4).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83246570; PubMed=6306594;
RA Fujiyama A., Miyazonara A., Nozaki C., Yoneyama T., Ohtomo N.,
RA Matsubara K.;
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
adr.";
RL Nucleic Acids Res. 11:4601-4610(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA](N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01587; CAA25746.1; -.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR004177; RYase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; ryl; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
Hydrolase; Nuclease; Endonuclease; DNA replication; DNA-binding.
KW SEQUENCE 730 AA; 82207 MW; 898A8BB50D96F2 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 730;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSRGSGS 19
DB 233 ARGSRGSGS 242

RESULT 9
DPOI_HPBV4 STANDARD; PRT; 842 AA.
ID DPOI_HPBV4
AC P31870; Q67861; Q67866;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
```

```

DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P
OS Hepatitis B virus (subtype adr / mutant).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31512;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89183619; PubMed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
RT virus subtype adr."
RL Nucleic Acids Res. 17:2124-2124(1989).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14193; CAA32399.1; -.
DR EMBL; X14193; CAA32405.1; ALT_TERM.
DR PIR; S04568; JDIVSV.
DR InterPro; IPR001462; DNAPOL_viral_C.
DR InterPro; IPR000201; DNAPOL_viral_N.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvc; 2.
DR ProDom; PD000814; DNAPOL_viral_C; 1.
KW Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 842 AA; 94545 MW; 137FE37246BA5C6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 842;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242
:|:|:|:|:|:|
:|:|:|:|:|:|

RESULT 10
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID DPOL_HPBVR
AC P03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=306820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada K., Igatahi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adw."
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00867; -, NCT_ANNOTATED_CDS.
DR PIR; A00704; JDIVSV.
DR InterPro; IPR001462; DNAPOL_viral_C.
DR InterPro; IPR000201; DNAPOL_viral_N.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvc; 2.
DR ProDom; PD000814; DNAPOL_viral_C; 1.
KW Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 843 AA; 94400 MW; A6BD490839C438B CRC64;

Query Match 46.7%; Score 43; DB 1; Length 843;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242
:|:|:|:|:|:|
:|:|:|:|:|:|

RESULT 11
ROG_MOUSE STANDARD; PRT; 388 AA.
ID ROG_MOUSE
AC Q35479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein G (hnRNP G) (RNA binding motif
DE protein, X chromosome).
DE RNMK OR RNMXP1 OR RNMXP2 OR HNRPG OR HNRNP.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98119027; PubMed=9457683;
RA Delbridge M.L., Ma K., Subbarao M.N., Cooke H.J., Basasin S.,
RA Graves J.A.M.;
RT "Evolution of mammalian HNRPG and its relationship with the putative
RT azoospermia factor RBM."
RL Mamm. Genome 9:168-170(1998).
CC -1- FUNCTION: Binds RNA.
CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes.
CC -1- PTM: O-glycosylated (by similarity).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031568; AAB86639.1; -.
DR HSSP; P09651; 1HA1.
DR MG: MGI:1343045; Rbmxt.
DR GO: GO:0005634; C:nucleus; IDA.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SMO0360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.

```


RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dlatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton S., Kettelman N., Madan A., Rodriguez Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in intracellular signal transduction by
 CC amplifying type I and type II IFN signaling. Phosphorylates the
 CC interferon-alpha/beta receptor alpha chain. Plays an essential
 CC role in promoting selective immune responses, including innate
 CC host defense mechanisms and specific antiviral activities.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC MISCELLANEOUS: Tyk2 deficient mice display reduced responses to
 CC interferon-alpha/beta and IL-12 and a selective deficiency in
 CC STAT3 activation in these pathways. Interferon-gamma signaling is
 CC also impaired.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. JAK
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 FERM domain.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF173032; AAD49423.1; -;
 DR EMBL: AF052607; AAC34580.2; -;
 DR EMBL: BC019789; AAI19789.1; -;
 DR HSSP: P11362; IFGK.
 DR MGI: 1929470; Tyk2.
 DR InterPro: IPR000299; Band 4.1.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR008266; Tyr_pkinase_AS.
 DR Pfam: PF00069; Pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 2.
 DR SMART: SM00295; B4; 1.
 DR SMART: SM00220; S_TKc; 2.
 DR SMART: SM00219; TYRK; 2.
 DR PROSITE: PS00660; FERM_1; FALSE_NEG.
 DR PROSITE: PS00661; FERM_2; FALSE_NEG.
 DR PROSITE: PS00577; FERM_3; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_DOM; 2.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00001; SH2; FALSE_NEG.
 DR Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW SH2 domain; Repeat.
 FT DOMAIN 29 426 FERM.
 FT DOMAIN 445 525 SH2 (ATYPICAL).
 FT DOMAIN 585 862 PROTEIN KINASE 1.
 FT DOMAIN 890 1162 PROTEIN KINASE 2.
 FT NP_BIND 896 904 ATP (BY SIMILARITY).
 FT BINDING 923 923 ATP (BY SIMILARITY).

FT ACT_SITE 1016 1016 BY SIMILARITY.
 FT MOD_RES 1047 1047 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 191 191 C -> F (IN REF. 1; AAC34580).
 FT CONFLICT 234 234 H -> R (IN REF. 1; AAC34580).
 FT CONFLICT 355 355 K -> E (IN REF. 1; AAC34580).
 FT CONFLICT 831 831 C -> S (IN REF. 1).
 FT CONFLICT 1099 1099 T -> M (IN REF. 1).
 SQ SEQUENCE 1180 AA; 132857 MW; 65E060079F08B1C9 CRC64;
 Query Match 46.2%; Score 42.5; DB 1; Length 1180;
 Best Local Similarity 57.9%; Pred. No. 53;
 Matches 11; Conservative 1; Mismatches 2; Indels 5; Gaps 1;
 CQ 3 QESTRGXRG-----RGR 16
 Db 331 QESTRGSRGNPHGRSGK 349
 RESULT 14
 FRZ2 DROME STANDARD; PRT; 694 AA.
 AC Q9VYX3; Q94916; Q9VYX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frlizzled protein 2 precursor (Frlizzled-2) (Dfz2).
 OS Frlizzled protein 2 precursor (Frlizzled-2) (Dfz2).
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhanoor P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.;
 RT "A new member of the frizzled family from Drosophila functions as a
 RT WINGLESS receptor."
 RL Nature 382:225-230(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beckley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Peiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolashkov S.,
 RA Borokov D., Borhan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Butts K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport U.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.O., Wei M.-H., Ibegwam C.,
 RA Jaitai B., Kalush F., Karpen G.H., Ke Z., Kennison U.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasbo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nusser D.R., Padgett J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard V., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RL "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 are coupled to the beta-catenin canonical signaling pathway, which
 leads to the activation of dishevelled proteins. Inhibition of
 CC GSX-3 kinase, nuclear accumulation of beta-catenin and activation
 of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeleton of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage
 CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U65569; AAC47273.1; -
 DR EMBL, AE003518; AA69185.2; -
 DR PIR, S71786; S71786;
 DR FlyBase, FBgn0016797; fz2.
 DR GO, GO:0016021; C:Integral to membrane; NMS.
 DR GO, GO:0017147; F:Wnt-protein binding; IDA.
 DR GO, GO:0006928; P:cell motility; IMP.
 DR GO, GO:0007163; P:establishment and/or maintenance of cell po. .; ISS.
 DR GO, GO:0008585; P:female gonad development; IMP.
 DR GO, GO:0016055; P:Wnt receptor signaling pathway; IDA.
 DR InterPro, IPR000539; Frizzled.
 DR InterPro, IPR000832; GPCR secretin.
 DR InterPro, IPR01534; GPCR secretin.
 DR Pfam, PF01392; Frizzled_1.
 DR Pfam, PF01392; Fz_1.
 DR PRINTS, PR00489; FRIZZLED.
 DR SMART, SM00063; FRI1.1.
 DR PROSITE, PS50038; Fz_1.
 DR PROSITE, PS50261; G-PROTEIN RECP F2_4; 1.
 KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
 KW Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 694
 FT DOMAIN 23 315
 FT TRANSMEM 316 336
 FT 1 (POTENTIAL).

FT DOMAIN 337 352
 FT TRANSMEM 353 373
 FT DOMAIN 374 397
 FT TRANSMEM 398 418
 FT TRANSMEM 419 439
 FT TRANSMEM 440 460
 FT TRANSMEM 461 482
 FT TRANSMEM 483 503
 FT TRANSMEM 504 524
 FT TRANSMEM 535 555
 FT TRANSMEM 556 584
 FT TRANSMEM 585 605
 FT TRANSMEM 606 694
 FT DOMAIN 187 225
 FT DOMAIN 59 180
 FT SITE 608 613
 FT SITE 692 694
 FT CARBOHYD 78 78
 FT CARBOHYD 288 288
 FT CONFLICT 55 55
 FT CONFLICT 417 417
 SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAB0956 CRC64;
 Query Match 45.7%; Score 42; DB 1; Length 694;
 Best Local Similarity 52.6%; Pred. No. 35;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHOESTRGSRGSRGSGS 19
 DB 182 SYTEAGSGSGSGSGSGS 200
 RESULT 15
 RN1_FUGRU STANDARD; PRT; 1097 AA.
 ID RN1_FUGRU
 AC Q98TR3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative regulator of nonsense transcripts 1.
 GN RENT1.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCB1_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clarke D., Elgar G., Clark M.S.;
 RT "Comparative analysis of human 19p12-13 region in Fugu and mouse.";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Eliminates the production of nonsense-containing RNAs
 CC (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AJ301641; CAC33025.1; -
 KW Nonsense-mediated mRNA decay; Hydrolyase; Helicase; ATP-binding;
 KW Zinc-finger.
 FT ZN FING 113 141
 FT ZN FING 165 195
 FT NP BIND 474 481
 FT NP BIND 1097 AA; 122338 MW; FDC7C2B4E26AC54 CRC64;
 SQ SEQUENCE

Tue Sep 28 07:11:28 2004

us-09-308-150-2.open.rsp

Page 9

Query Match	45.7%;	Score 42;	DB 1;	Length 1097;
Best Local Similarity	41.7%;	Pred. No. 58;		
Matches 10;	Conservative 3;	Mismatches 5;	Indels 6;	Gaps 1;

```
QY      1 SHQEST-----RGXSRGRSGRSG 18
          |||         |||
Db       975 SYQGQTNGPAAGRGAMKGKSGRGG 998
```

Search completed: September 28, 2004, 06:05:46
Job time : 8.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-2

Perfect score: 92
Sequence: 1 SHOESTRGSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_Organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	797	4	Q16824
2	90	97.8	590	4	Q15206
3	90	97.8	1218	4	Q05331
4	84	91.3	798	4	Q09403
5	84	91.3	1084	4	Q01212
6	75	81.5	591	4	Q01720
7	75	81.5	687	4	Q09402
8	73	79.3	322	4	Q075370
9	70	76.1	465	4	Q03838
10	50	54.3	800	5	Q15845
11	49	53.3	213	15	Q07818
12	48	52.2	327	12	Q09331
13	48	52.2	700	13	Q042378
14	48	52.2	715	13	Q080F00
15	48	52.2	715	13	Q080G68
16	48	52.2	716	13	Q042107

17	48	52.2	822	3	Q09312	Q09312 neurospora
18	48	52.2	1229	3	Q062H1	Q062H1 neurospora
19	47	51.1	471	5	Q09AY0	Q09AY0 caenorhabdi
20	46	50.0	197	5	Q17626	Q17626 caenorhabdi
21	46	50.0	203	5	P01026	P01026 caenorhabdi
22	46	50.0	236	10	Q09AY4	Q09AY4 oryza sativ
23	46	50.0	236	10	Q07X25	Q07X25 oryza sativ
24	46	50.0	353	10	Q085W9	Q085W9 oryza sativ
25	46	50.0	399	4	Q0NTA9	Q0NTA9 homo sapien
26	46	50.0	467	4	Q0BWT8	Q0BWT8 homo sapien
27	46	50.0	737	4	Q0B039	Q0B039 homo sapien
28	46	50.0	836	10	Q0SKZ9	Q0SKZ9 arabidopsis
29	46	50.0	843	12	Q0QNS2	Q0QNS2 hepatitis b
30	46	50.0	1378	5	Q07405	Q07405 halictis di
31	45	48.9	96	15	Q0JB05	Q0JB05 human immun
32	45	48.9	96	15	Q0WRP4	Q0WRP4 human immun
33	45	48.9	96	15	Q0JBV4	Q0JBV4 human immun
34	45	48.9	96	15	Q0ADP5	Q0ADP5 human immun
35	45	48.9	222	4	Q06E40	Q06E40 homo sapien
36	45	48.9	243	10	Q0AX45	Q0AX45 oryza sativ
37	45	48.9	349	16	Q07U67	Q07U67 rhodospirill
38	45	48.9	374	5	Q07651	Q07651 caenorhabdi
39	45	48.9	476	12	Q080S90	Q080S90 herpesvirus
40	45	48.9	652	5	Q07168	Q07168 apis mellif
41	45	48.9	722	13	Q080M4	Q080M4 brachydantio
42	45	48.9	1096	10	Q04158	Q04158 oryza sativ
43	45	48.9	1275	4	Q0UQ36	Q0UQ36 homo sapien
44	45	48.9	1754	11	Q080U37	Q080U37 mus musculu
45	45	48.9	1783	4	Q15038	Q15038 homo sapien

ALIGNMENTS

RESULT 1
ID Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-UN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride C.W., Idler W.W., Markova N., Stelvert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene [published erratum appears in Biochemistry 1991 Jun 11;30(23):5814]."
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60502; AAA63248.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03515; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDA86B CRC64;
Query Match 97.8%; Score 90; DB 4; Length 797;
Best Local Similarity 94.7%; Pred. No. 2,3e+06;
Matches 18; Conservative 0; Mismatches 1; Indels 0;
Gy 1 SHOESTRGSRGRSGSGS 19
Db 427 SHOESTRGSRGRSGSGS 445

RESULT 2

Q15206 ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene [published erratum appears in Biochemistry 1991 Jun 11;30(23):5814]."
RT Biochemistry 29:9432-9440(1990).
RL Biochemistry 29:9432-9440(1990).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene."
RT Biochemistry 30:5814-5814(1991).
RL EMBL; M60494; AAA63244.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON_TER 990 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
Query Match 97.8%; Score 90; DB 4; Length 990;
Best Local Similarity 94.7%; Pred. No. 2.3e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 1 SH0ESTRGXSRGSRGSGS 19
Db 551 SH0ESTRGXSRGSRGSGS 569
RESULT 3
Q05331 ID Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109349; PubMed=8417356;
RA Markova N.G., Markov L.N., Chipev C.C., Gan S.-O., Idler W.W., Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein."
RT Mol. Cell. Biol. 13:613-625(1993).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.
CC EMBL; M96943; AAA36487.1; -
CC PIR; A48118; A48118.
CC HSPB; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; F:cell differentiation; NAS.
DR GO; GO:0008151; F:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_5100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eefhand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding; KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;
Query Match 97.8%; Score 90; DB 4; Length 1218;
Best Local Similarity 94.7%; Pred. No. 3.7e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 1 SH0ESTRGXSRGSRGSGS 19
Db 773 SH0ESTRGXSRGSRGSGS 791
RESULT 4
Q094U3 ID Q094U3 PRELIMINARY; PRT; 798 AA.
AC Q094U3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE D14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RA Laitid G.;
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL356504; CAC13171.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; Filaggrin.
DR NON_TER 1 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 2.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 SH0ESTRGXSRGSRGSGS 19
Db 428 SH0ESTRGXSRGSRGSGS 446
RESULT 5
Q01212 ID Q01212 PRELIMINARY; PRT; 1084 AA.

```

AC 001212; 003840;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-2003 (TEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 91.3%; Score 84; DB 4; Length 1084;
Best Local Similarity 89.5%; Pred. No. 3.2e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 SHOESTRCKSRGSRGSGS 19
DB 64 SHOESTRCKSRGSRGSGS 82

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; IEBA.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467
FT PROPEP 468 474
FT CHAIN 475 >591
FT CA_BIND 19 32
FT CA_BIND 62 73
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.00051;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHOESTRCKSRGSRGSGS 19
DB 449 SHOESTRCKSRGSRGSGS 467

RESULT 7
Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE D1J4N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; IEBA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON TER 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 687;
Best Local Similarity 84.2%; Pred. No. 0.0006;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 SHOESTRGXSRGSRGSGS 19
 DB 449 SHOESTRGXSRGSRGSGS 467

RESULT 8

ID 075370 PRELIMINARY; PRT; 322 AA.

AC 075370;
 DT 01-NOV-1998 (TRMBLrel. 08, Created)
 DT 01-NOV-1998 (TRMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TRMBLrel. 24, Last annotation update)
 DE Epidermal filaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;
 RA Girbal-Neuhausser E., Durrieux J.U., Arnaud M., Dalbon P., Sebbag M.,
 RA Vincent C., Simon M., Sennhu T., Maeson-Bessiere C.,
 RA Jolivet-Reynaud C., Jolivet M., Serre G.;
 RT "The epitopes targeted by the rheumatoid arthritis-associated
 RT anti-filaggrin autoantibodies are posttranslationally generated on
 RT various sites of (pro)filaggrin by deamination of arginine residues."
 RL J. Immunol. 162:585-594(1999).
 DR EMBL; AF043380; AAC23559.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 2.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON_TER 1
 FT NON_TER 322
 SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230DBFF9E0 CRC64;

Query Match 79.3%; Score 73; DB 4; Length 322;
 Best Local Similarity 83.3%; Pred. No. 0.00057;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRGSGS 18

DB 305 SHOEFTRGSGRSGSGS 322

RESULT 9

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003838;
 DT 01-NOV-1996 (TRMBLrel. 01, Created)
 DT 01-NOV-1996 (TRMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
 DE FILAGGRIN (PROFILAGGRIN) (Fragment).
 GN FLG.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 RT gene."
 RL Biochemistry 29:9432-9440(1990).
 RN [2]
 RP REVISIONS
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 RT gene."
 RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
 CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
 CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 CC -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
 CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
 CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN
 CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 DR EMBL; M60499; AAA63246.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 3.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON_TER 1
 FT NON_TER 465
 SQ SEQUENCE 465 AA; 50280 MW; C683744C5E134097 CRC64;

Query Match 76.1%; Score 70; DB 4; Length 465;
 Best Local Similarity 78.9%; Pred. No. 0.0027;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRGSGS 19

DB 227 SHOESTRGXSRGSRGSGS 245

RESULT 10

ID 015845 PRELIMINARY; PRT; 800 AA.

AC 015845;
 DT 01-JAN-1998 (TRMBLrel. 05, Created)
 DT 01-JAN-1998 (TRMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TRMBLrel. 24, Last annotation update)
 DE L1439.4.
 GN L1439.4.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5664;
 RN [1]

RP SEQUENCE FROM N.A.
 RX STRAIN=Frledlin;
 RX MEDLINE=99178987; PubMed=10077609;
 RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
 RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
 RA Bastien P., Fu G., Ivans A., Stuart K.;
 RT "Leishmania major Frledlin chromosome 1 has an unusual distribution of
 RT protein-coding genes."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
 DR EMBL; AB001274; AAC24675.1; -;
 DR PIR; T02852; T02852.
 SQ SEQUENCE 800 AA; 83782 MW; DA57AB627913CBAC CRC64;

Query Match 54.3%; Score 50; DB 5; Length 800;
 Best Local Similarity 55.6%; Pred. No. 9.7;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTRGXSRGSRGSGS 19

DB 429 HRDGVGRGJSTASGRSGS 446

RESULT 11

ID 072B18 PRELIMINARY; PRT; 213 AA.

AC 072B18;
 DT 01-JUN-2003 (TRMBLrel. 24, Created)
 DT 01-JUN-2003 (TRMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
 DE Vif protein.
 GN VIF.

OS Chimpanzee immunodeficiency virus (SIV(cp2)) (CIV).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxId=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SIMmd5440;
 RX MEDLINE=22550088; PubMed=12663793;
 RA Hu J., Switzer W.M., Foley B.T., Robertson D.L., Goeken R.M.,
 Korber B.T., Hirsch V.M., Beer B.E.,
 RT "Characterization and Comparison of Recombinant Simian
 Immunodeficiency Virus from Drill (Mandrillus leucophaeus) and
 RT Mandrill (Mandrillus sphinx) Isolates."
 RL J. Virol. 77:4867-4880(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SIMmd5440;
 RA Hu J., Hirsch V.M., Beer B.E.,
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY159322; AAC22476.1; -.
 DR GO; GO:0019058; P:Viral infectious cycle; IEA.
 DR InterPro; IPR000475; Viral_infect.
 DR Pfam; PF00559; Vif.1
 DR PRINTS; PR00349; VIRIONINPCT.
 DR PRODOM; PD000063; Viral_infect.1.
 SQ SEQUENCE 213 AA; 24885 MW; EA626C2C0BP4ED27 CRC64;

Query Match 53.3%; Score 49; DB 15; Length 213;
 Best Local Similarity 52.9%; Pred. No. 3 4;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGSGRSGRSGS 19
 Db 163 REETRGARRSRGRGTGA 179

RESULT 12
 Q993H1 PRELIMINARY; PRT; 327 AA.
 ID Q993H1;
 AC Q993H1;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE ORF3.
 OS Callitriche herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxId=106331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CJ0149;
 RX MEDLINE=21107697; PubMed=11158621;
 RA Wang F., Ramer J., Rivallier P., Quink C., Garber R.L., Beier D.R.,
 RT "An Epstein-Barr-related herpesvirus from marmoset lymphomas."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:1224-1228(2001).
 DR EMBL; AF319782; AKK38247.1; -.
 DR HSP; P03211; IYH1.
 DR GO; GO:0042025; C:host cell nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR004186; EBNA1.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF02905; EBNA1; 1.
 SQ SEQUENCE 327 AA; 34334 MW; 79F03C1D2015C94A CRC64;

Query Match 52.2%; Score 48; DB 12; Length 327;
 Best Local Similarity 75.0%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RGSRGSGRSG 18
 Db 69 RGSRGSGRSG 80

RESULT 13

O42378 PRELIMINARY; PRT; 700 AA.
 ID O42378;
 AC O42378;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE RNA helicase (DEAD box).
 GN VASA OR VLG.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98025484; PubMed=9376327;
 RA Olsen L.C., Aasland R., Fjose A.;
 RT "A vasa-like gene in Zebrafish identifies putative primordial germ
 RT cells."
 RL Mech. Dev. 66:95-105(1997).
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 DR EMBL; Y12007; CAAT2735.1; -.
 DR HSP; Q58083; IYH8.
 DR ZFIN; ZDB-GENE-990415-272; vasa.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0003476; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR00629; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 DR ATP-binding; Helicase; Hydrolase.
 KW SEQUENCE 700 AA; 75312 MW; 726B56A90DEC2D83 CRC64;

Query Match 52.2%; Score 48; DB 13; Length 700;
 Best Local Similarity 44.4%; Pred. No. 18;
 Matches 12; Conservative 1; Mismatches 6; Indels 8; Gaps 1;

QY 1 SHQSTRGXS-----RGSRGSGS 19
 Db 26 SNSEGTGSSWKMGTGDSFRGSGRSGS 52

RESULT 14
 Q8QF00 PRELIMINARY; PRT; 715 AA.
 ID Q8QF00;
 AC Q8QF00;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Vase-like protein.
 GN VLG.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vatland A., Olsen L.C.;
 RT "Establishment of transgenic zebrafish lines expressing green
 RT fluorescent protein in their germ cells."
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 DR EMBL; AJ311625; CAC84069.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 / Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-2

Perfect score: 92
Sequence: 1 SHQESTRXSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseg_29Jan04:*
1: genesegp1980s:*
2: genesegp1990s:*
3: genesegp2000s:*
4: genesegp2001s:*
5: genesegp2002s:*
6: genesegp2003ae:*
7: genesegp2003bs:*
8: genesegp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	90	97.8	19	2	AAW61506 Peptide c
2	90	97.8	19	2	AAW61514 Peptide c
3	90	97.8	1467	5	ABH97605 Novel hum
4	86	93.5	19	2	AAW61517 Peptide c
5	85	92.4	19	2	AAW61516 Peptide c
6	84	91.3	19	2	AAW61508 Peptide c
7	84	91.3	19	2	AAW61515 Peptide c
8	84	91.3	19	2	AAW61507 Peptide c
9	84	91.3	19	2	AAW61509 Peptide c
10	84	91.3	19	2	AAW61510 Peptide c
11	84	91.3	19	2	AAW61505 Peptide c
12	78	84.8	19	2	AAW61512 Peptide c
13	78	84.8	19	2	AAW61511 Peptide c
14	78	84.8	19	2	AAW61513 Peptide c
15	75	81.5	21	2	AAW61520 Peptide c
16	75	81.5	22	4	AAW61523 IGP1546 P
17	59	64.1	330	2	AAW22956 Human fil
18	59	64.1	330	2	AAW22955 Human fil
19	59	64.1	330	2	AAW22957 Human fil
20	55	59.8	330	2	AAW22954 Human fil
21	51	55.4	1711	4	AAW79819 Human pro
22	51	55.4	1951	4	AAW78835 Human pro
23	50	54.3	477	6	ABO07142 Novel hum
24	50	54.3	2348	4	ABG10929 Novel hum
25	48	52.2	700	4	AAW02421 Danio rei

26	47	51.1	184	3	AAB57041 Human pro
27	47	51.1	1038	7	ADC03412 Rice flow
28	46	50.0	443	4	AAW40069 Human pol
29	46	50.0	574	4	AAW41855 Human pol
30	46	50.0	752	4	AAW23535 Novel hum
31	46	50.0	19938	6	ABP76682 Streptomy
32	45	48.9	222	4	AAW93652 Human pol
33	45	48.9	1791	6	ABO52986 Human put
34	44	47.8	56	4	AAW60540 Propionib
35	44	47.8	56	6	ABW57059 Propionib
36	44	47.8	134	4	ABG23065 Novel hum
37	44	47.8	157	3	ABG41022 Zea mays
38	44	47.8	173	6	ABR41796 Human DIT
39	44	47.8	215	3	AAW46917 Arabidops
40	44	47.8	229	7	ADW59825 Rat Prote
41	44	47.8	229	7	ADW59833 Rat Prote
42	44	47.8	229	7	ADW59837 Rat Prote
43	44	47.8	229	7	ADW59829 Rat Prote
44	44	47.8	258	3	AAW46916 Arabidops
45	44	47.8	297	3	AAW46915 Arabidops

ALIGNMENTS

RESULT 1
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX
AC AAW61506;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 9 /note= "Citruiline"
FT XX
FN W09822503-A2.
XX
PD 28-MAY-1998.
XX
PD 14-NOV-1997; 97WO-NL000624.
XX
PF 15-NOV-1996; 96NL-01004539.
XX
PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PR (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PA Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
XX
PI WPT, 1998-398613/34.
XX
DR Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
PS Disclosure; Page 6, 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;
 SQ

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 1 SHQESTRGXSRGSRGSGS 19

RESULT 2

ID AAW61514 standard; peptide; 19 AA.

XX AAW61514;

XX 26-OCT-1998 (first entry)

XX Peptide cf, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPT; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 1.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 1 SHQESTRGXSRGSRGSGS 19

ABB97605
 ID ABB97605 standard; protein; 1467 AA.
 XX ABB97605;
 AC

XX 27-UN-2002 (first entry)

XX Novel human protein SEQ ID NO: 873.

XX Human; antihaemic; vulnary; antiinflammatory; immunomodulator;

XX antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;

XX neuropeptide; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

XX Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI: 2002-292408/33.

XX N-PSDB; ABB32791.

XX An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

XX novel human proteins. These were isolated from expressed sequences tags

XX (ESTs). They can be used to stimulate cell growth, to regulate

XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

XX e.g. in burn treatment, to regulate the immune system e.g. to treat

XX multiple sclerosis, to regulate activin or inhibin e.g. to treat

XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.

XX rheumatoid arthritis, and to treat nervous system disorders e.g.

XX Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.4e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 773 SHQESTRGXSRGSRGSGS 791

RESULT 4
 ID AAW61517
 XX AAW61517 standard; peptide; 19 AA.
 AC AAW61517;
 XX 26-OCT-1998 (first entry)
 XX Peptide cfo, based on cDNA of a profilaggrin repeat.
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

RESULT 3

```

OS Synthetic.
OS Homo sapiens.
XX WO3822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match: 93.5%; Score 86; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred.No. 7.8e-07;
XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0
XX
XX QY 1 SHOESTRGXSRGSRGSGS 19
XX |||||:|||||||
XX 1 SHOESTGRSGRSGRSGSGS 19
XX
XX Db 1 SHOESTGRSGRSGRSGSGS 19
XX
XX RESULT 5
XX AAW61516
XX ID AAW61516 standard; peptide; 19 AA.
XX
XX AC AAW61516;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE Peptide cIE, based on cDNA of a profilaggrin repeat.
XX
XX KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX KW monoclonal antibody.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX PN WO3822503-A2.
XX
XX PD 28-MAY-1998.
XX
XX PF 14-NOV-1997; 97WO-NL000624.
XX
XX PR 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

```

XX	WP1, 1998-398613/34.
DR	
XX	Peptide derived from an antigen recognised by autoantibodies - is
PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT	used in diagnosis of the disease.
XX	
XX	Disclosure; Page 6; 19pp; English.
PS	
CC	Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC	the profilaggrin antigen which is recognised by autoantibodies from
CC	patients with Rheumatoid arthritis (RA). This peptide is reactive with a
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC	The peptides were created by using standard solid phase synthesis, which
CC	produced them as peptide amides. These sequences may be used in the
CC	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC	for obtaining polyclonal and monoclonal antibodies
XX	
SQ	Sequence 19 AA;
XX	
Query Match	92.4%; Score 85; DB 2; Length 19;
Best Local Similarity	89.5%; Pred. No. 1.1e-06;
Matches 17; Conservative	0; Mismatches 2; Indels 0; Gaps 0
QY	1 SHOESTRGXSRGRSGSGS 19
D6	1 SHOESTRGSRGRSGSGS 19
XX	
RESULT 6	
AAM61508	
ID	AAM61508 standard; peptide, 19 AA.
XX	
AC	AAM61508;
XX	
DT	26-OCT-1998 (first entry)
XX	
DE	Peptide cfcd4, based on cDNA of a profilaggrin repeat.
XX	
KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW	solid phase synthesis; peptide amide; polyclonal antibody;
XX	monoclonal antibody.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
XX	
FT	Key Location/Qualifiers
FT	Modified-site 13
FT	/note= "Citruilins"
XX	
FN	W09822503-A2.
XX	
PD	28-MAY-1998.
XX	
PF	14-NOV-1997; 97MO-NL000624.
XX	
PR	15-NOV-1996; 96NL-01004539.
XX	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.
PI	Van Venrooij MJW, Schellekens GA, Raats JWH, Hoet RMA;
XX	
DR	WP1; 1998-398613/34.
XX	
PT	Peptide derived from an antigen recognised by autoantibodies - is
PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT	used in diagnosis of the disease.
XX	
PS	Disclosure; Page 6; 19pp; English.
XX	
CC	Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC	the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SO Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGXSGRSGSGS 19
 |||||
 Db 1 SHOESTRGXSGRSGSGS 19

RESULT 7
 AAW61515
 ID AAW61515 standard; peptide; 19 AA.

AC AAW61515;

DT 26-OCT-1998 (first entry)

DE Peptide cfa, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

OS Synthetic.

OS Homo sapiens.

PN WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGXSGRSGSGS 19
 |||||

Db 1 SHOESTAGSRGSRSGS 19

RESULT 8
 AAW61507
 ID AAW61507 standard; peptide; 19 AA.

AC AAW61507;

DT 26-OCT-1998 (first entry)

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

OS Synthetic.

OS Homo sapiens.

PN WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGXSGRSGSGS 19
 |||||
 Db 1 SHOESTRGXSGRSGSGS 19

RESULT 9

ID AAW61509 standard; peptide; 19 AA.

AC AAW61509;

DT 26-OCT-1998 (first entry)

XX

DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 16
 FT /note= "Citruiline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 XX used in diagnosis of the disease.

XX PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

XX Query Match 91.3%; Score 84; DB 2; Length 19;
 XX Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
 DB 1 SHQESTRGXSRGRSGSGS 19

RESULT 10
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

XX 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruiline"
 FT Modified-site 9
 FT /note= "Citruiline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 XX used in diagnosis of the disease.

XX PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

XX Query Match 91.3%; Score 84; DB 2; Length 19;
 XX Best Local Similarity 94.7%; Pred. No. 1.6e-06;
 XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
 DB 1 SHQESTRGXSRGRSGSGS 19

RESULT 11
 AAW61505
 ID AAW61505 standard; peptide; 19 AA.

XX AAW61505;

XX 26-OCT-1998 (first entry)

XX Peptide cfc1, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruiline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
 Db 1 SHQESTXGRSGRSGSGS 19

RESULT 14
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.

AC AAW61513;
 DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 7
 FT Modified-site /note= "Citruilline"
 FT Modified-site 16
 FT Modified-site /note= "Citruilline"

PN W09822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
 Db 1 SHQESTXGRSGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 DT 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 9
 FT Modified-site /note= "Citruilline"

PN W09822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 4.9e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HQESTRGXSRGRSGSGS 19
 Db 4 HQESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:11
 Job time : 56.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHQSTRGKSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-2
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-3
7	84	91.3	19	9	US-09-308-150-4
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-6
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-7
12	78	84.8	19	9	US-09-308-150-8
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	50	54.3	477	15	US-10-161-927-62	Sequence 62, Appl
17	48	52.2	506	16	US-10-437-963-185974	Sequence 185974,
18	47	51.1	184	9	US-09-925-100-1619	Sequence 1619, Ap
19	47	51.1	257	16	US-10-437-963-105967	Sequence 105967,
20	46.5	50.5	176	16	US-10-437-963-196023	Sequence 196023,
21	46	50.0	16	10	US-09-791-524-150	Sequence 150, App
22	46	50.0	130	16	US-10-437-963-104588	Sequence 104588,
23	46	50.0	224	12	US-10-424-599-269475	Sequence 269475,
24	46	50.0	226	16	US-10-437-963-102742	Sequence 102742,
25	46	50.0	445	12	US-10-424-599-267922	Sequence 267922,
26	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
27	46	50.0	1087	16	US-10-641-991-24	Sequence 8, Appl
28	46	50.0	19608	15	US-10-084-846A-8	Sequence 30685, A
29	45	48.9	79	14	US-10-029-366-103865	Sequence 103033,
30	45	48.9	213	16	US-10-437-963-103033	Sequence 1630387,
31	45	48.9	294	16	US-10-437-963-169387	Sequence 169387,
32	45	48.9	563	16	US-10-437-963-198539	Sequence 198539,
33	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
34	45	48.9	2263	16	US-10-408-765A-2231	Sequence 2231, Ap
35	44.5	48.4	307	12	US-10-425-114-39059	Sequence 39059, A
36	44	47.8	78	16	US-10-437-963-203073	Sequence 203073,
37	44	47.8	430	16	US-10-437-963-164039	Sequence 164039,
38	44	47.8	511	16	US-10-437-963-188040	Sequence 188040,
39	44	47.8	1071	16	US-10-437-963-118389	Sequence 118389,
40	43.5	47.3	85	12	US-10-424-599-161928	Sequence 161928,
41	43.5	47.3	181	16	US-10-437-963-117640	Sequence 117640,
42	43.5	47.3	209	16	US-10-437-963-175682	Sequence 175682,
43	43	46.7	50	14	US-10-029-386-27692	Sequence 27692, A
44	43	46.7	76	12	US-10-424-599-248912	Sequence 248912,
45	43	46.7	87	9	US-09-867-550-764	Sequence 764, App

ALIGNMENTS

RESULT 1
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Walthers Jacobus Wilhelmus
APPLICANT: Scheilekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoel, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.86-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 2

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.8e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 3

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 19

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 5

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-1

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2,1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 6
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-3

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2,1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 7
US-09-308-150-4

Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-4

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2,1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 8
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-5

US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 9

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.1e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 10

US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 11

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 12

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.0015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

```

```

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137032A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.0015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00045;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 HOESTXGRSGRSGSGS 19
Db 4 HOESTXGRSGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NPU500 INNS:031
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US/09/747,029A
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```

```

: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic Peptides
: NAME/KEY: MOD.RES
: LOCATION: (1)..(10)
: OTHER INFORMATION: Xaa represents Citrulline
TS-09-747-029A-22

```

Query Match	81.5%	Score 75;	DB 9;	Length 22;
Best Local Similarity	83.3%;	Pred. No. 0.00047;		
Matches 15; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

```
Qy      2 HQESTRGXSRGRSGS 19
          |||||
Db      5 HQESTXGRRSRGRSGS 22
```

```
Search completed: September 28, 2004, 07:28:49
Job time : 111.526 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHOESTRGRSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	51.1	406	4	US-09-252-991A-1857
2	47	51.1	406	4	US-09-252-991A-1857
3	46	50.0	465	4	US-09-252-991A-18919
4	45	48.9	204	4	US-09-252-991A-21317
5	44	47.8	369	4	US-09-252-991A-20245
6	44	47.8	483	4	US-09-252-991A-19015
7	44	47.8	566	4	US-09-252-991A-24514
8	43	46.7	201	4	US-09-252-991A-25745
9	43	46.7	508	4	US-09-252-991A-28613
10	43	46.7	653	4	US-09-513-057C-13
11	43	46.7	720	4	US-08-591-502B-49
12	43	46.7	842	4	US-08-591-502B-50
13	43	46.7	842	4	US-08-591-502B-51
14	43	46.7	843	4	US-09-719-528A-2
15	43	46.7	843	4	US-08-591-502B-45
16	43	46.7	843	4	US-08-591-502B-48
17	43	46.7	843	4	US-08-591-502B-59
18	43	46.7	845	4	US-08-591-502B-46
19	42	45.7	138	4	US-09-252-991A-16961
20	42	45.7	169	3	US-09-342-084-6
21	42	45.7	250	4	US-09-252-991A-26599
22	42	45.7	274	4	US-09-252-991A-16696
23	42	45.7	629	4	US-09-252-991A-22901
24	42	45.7	703	4	US-09-252-991A-20353
25	42	45.7	2237	1	US-08-354-973-1
26	41	44.6	143	3	US-08-868-699A-4
27	41	44.6	143	4	US-09-757-014-4

28	41	44.6	205	4	US-09-252-991A-18529	Sequence 16529, A
29	41	44.6	219	4	US-09-621-976-4221	Sequence 4221, Ap
30	41	44.6	255	4	US-09-252-991A-19598	Sequence 19598, A
31	41	44.6	286	4	US-09-252-991A-22717	Sequence 22717, A
32	41	44.6	354	4	US-09-252-991A-22717	Sequence 22717, A
33	41	44.6	404	4	US-09-489-039A-11606	Sequence 11606, A
34	41	44.6	409	4	US-09-489-039A-14217	Sequence 14217, A
35	41	44.6	446	4	US-09-252-991A-21791	Sequence 21791, A
36	41	44.6	539	4	US-09-252-991A-17070	Sequence 17070, A
37	41	44.6	575	3	US-08-922-865-2	Sequence 2, Appli
38	41	44.6	575	4	US-09-510-948-2	Sequence 2, Appli
39	41	44.6	722	4	US-09-984-890-4	Sequence 4, Appli
40	41	44.6	724	4	US-09-984-890-2	Sequence 2, Appli
41	41	44.6	745	4	US-09-523-849-36	Sequence 36, Appli
42	41	44.6	796	3	US-08-868-699A-2	Sequence 2, Appli
43	41	44.6	796	4	US-09-757-014-2	Sequence 2, Appli
44	41	44.6	803	4	US-09-252-991A-23614	Sequence 23614, A
45	40.5	44.0	255	3	US-09-342-084-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-19857
; Sequence 19857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19857
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19857

Query Match 51.1%, Score 47; DB 4; Length 406;
Best Local Similarity 50.0%; Pred. No. 8.7;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSGRSGRSGS 19
DB 127 HGQSRGDRAGRGHGA 144

RESULT 2
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326

Query Match          51.1%; Score 47; DB 4; Length 421;
Best Local Similarity 52.9%; Pred. No. 9.1;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGSRGSRGSG 18
    ||| ||| ||| |||
    ||| ||| ||| |||
Db 13 HARTGTGTGRGRGRSG 29

RESULT 3
US-09-252-991A-18919
; Sequence 18919, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18919
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18919

Query Match          50.0%; Score 46; DB 4; Length 465;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSGS 19
    ||| ||| ||| |||
    ||| ||| ||| |||
Db 83 SGQTCRGALRGSGGRAGS 101

RESULT 4
US-09-252-991A-21317
; Sequence 21317, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21317
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21317

Query Match          48.9%; Score 45; DB 4; Length 204;
Best Local Similarity 58.8%; Pred. No. 8.8;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTRGSRGSRGSRGSGS 19
    ||| ||| ||| |||
    ||| ||| ||| |||
Db 67 QESRDGRGRSTRDRSGS 83

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20245

RESULT 5
US-09-252-991A-20245
; Sequence 20245, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20245
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20245

Query Match          47.8%; Score 44; DB 4; Length 369;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGSRGSRGSG 18
    ||| ||| ||| |||
    ||| ||| ||| |||
Db 221 HRRRTGSAAGGGRHG 237

RESULT 6
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015

Query Match          47.8%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGSRGSRGSG 18
    ||| ||| ||| |||
    ||| ||| ||| |||
Db 94 HABQDPGAGTGRAGRAG 110

RESULT 7
US-09-252-991A-24514
; Sequence 24514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

Tue Sep 28 07:11:27 2004

us-09-308-150-2.open.ra1

Page 3

;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 24514
;; LENGTH: 586
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24514

Query Match 47.8%; Score 44; DB 4; Length 586;
Best Local Similarity 61.1%; Pred. No. 37;
Matches 11; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 3 QESTR--GKSRGSGSG 18
Db 245 QRSRRTGGGSGSGSG 262

RESULT 8
US-09-252-991A-25745
; Sequence 25745, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25745
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25745

Query Match 46.7%; Score 43; DB 4; Length 201;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 HOESTRGXSRGSGR 16
Db 111 HSRWRGAGRGGRPR 125

RESULT 9
US-09-252-991A-28613
; Sequence 28613, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28613
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28613

Query Match 46.7%; Score 43; DB 4; Length 508;
Best Local Similarity 44.4%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHQESTRGXSRGSGRSG 18
Db 106 ARSRBRGAGSRGRTGRG 123

RESULT 10
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 46.7%; Score 43; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 59;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SHQESTRGXSRGSGRSGS 19
Db 573 SRQVSTASASAGRGISGS 591

RESULT 11
US-08-591-502B-49
; Sequence 49, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08665
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Laver

REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-591-502B-49

Query Match 46.7%; Score 43; DB 4; Length 730;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGKSGRSGS 242

RESULT 12
US-08-591-502B-50
Sequence 50, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:

US-08-591-502B-50

Query Match 46.7%; Score 43; DB 4; Length 842;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGKSGRSGS 242

RESULT 13
US-08-591-502B-51
Sequence 51, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51

Query Match 46.7%; Score 43; DB 4; Length 842;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGKSGRSGS 242

RESULT 14
US-09-719-528A-2

Sequence 2, Application US/09719528A
Patent No. 6558675
GENERAL INFORMATION:
APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi
Chen, Wei Ning
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USERS THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Iadas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford C.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2
Query Match 46.7%; Score 43; DB 4; Length 843;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 10 SRGRSGRSGS 19
:|||||
Db 233 ARGKSGRSGS 242
RESULT 15
US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45
Query Match 46.7%; Score 43; DB 4; Length 843;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 10 SRGRSGRSGS 19
:|||||
Db 233 ARGKSGRSGS 242

Search completed: September 28, 2004, 06:26:41
Job time : 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds

(without alignments)
139,915 Million cell updates/sec

Title: US-09-308-150-3

Sequence: 1 SHOESTRGSRGSRGSRGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	416	2 A32947	filaggrin precursor
2	91	98.9	2248	2 A35938	profilaggrin - hum
3	82	89.1	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2ML47	E2 protein - human
5	54	58.7	471	2 T33997	hypothetical prote
6	50	54.3	123	2 T16234	hypothetical prote
7	50	54.3	800	2 T02852	probable membrane
8	48	52.2	822	2 T51049	related to nucleol
9	46	50.0	197	2 T18918	hypothetical prote
10	46	50.0	203	2 C87801	protein C10G11.9 l
11	46	50.0	306	2 T21220	hypothetical prote
12	46	50.0	373	2 T02976	probable DNA bindi
13	46	50.0	339	2 T46259	hypothetical prote
14	46	50.0	836	2 G84727	probable DNA topol
15	46	50.0	1105	2 T47582	hypothetical prote
16	45	48.9	374	2 T33328	hypothetical prote
17	45	48.9	1337	2 T30291	dextranase - strep
18	44	47.8	89	2 H87367	hypothetical prote
19	44	47.8	229	2 TC7219	nuclear protein SR
20	44	47.8	867	2 T27136	hypothetical prote
21	44	47.8	871	2 T27135	hypothetical prote
22	44	47.8	1496	2 T00499	probable retroelem
23	43.5	47.3	463	2 T51194	hypothetical prote
24	43	46.7	70	4 T52626	hypothetical FUS/E
25	43	46.7	286	4 S50855	neurotrophin-6 - s
26	43	46.7	462	2 T15543	hypothetical prote
27	43	46.7	462	2 S33798	FUS/CHOP mutant fu
28	43	46.7	526	1 S33799	RNA-binding protei
29	43	46.7	528	2 G02127	fus-like protein -

30	43	46.7	1232	2 S40766	hypothetical prote
31	43	46.7	1415	1 EDBEGA	immediate-early pr
32	43	46.7	1507	2 B47328	natural killer cel
33	42	45.7	150	2 T84778	very hypothetical
34	42	45.7	203	2 T25916	hypothetical prote
35	42	45.7	649	2 C71283	probable ATP-depen
36	42	45.7	694	2 S71786	wingless receptor
37	42	45.7	849	2 A96592	hypothetical prote
38	42	45.7	1015	2 T42013	frequency clock pr
39	42	45.7	1287	2 S55954	viral mRNA transla
40	42	45.7	2233	2 S63347	acetyl-CoA carboxy
41	41.5	45.1	79	2 A40459	nuclear phosphopro
42	41	44.6	350	2 A44843	EMW2p30-47 antigen
43	41	44.6	136	2 T35632	probable transpos
44	41	44.6	250	2 T29344	hypothetical prote
45	41	44.6	311	2 T45683	hypothetical prote

ALIGNMENTS

```
RESULT 1
A32947
filaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, D.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A>Title: Characterization of a cDNA clone encoding human filaggrin and localization of th
A:Reference number: A32947; MUID:89296901; PMID:2740331
A:Accession: A32947
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M24355; NID:G182604; PID:AAA52454.1; PID:G182605
A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: Ep hand; epidermis; polymorphism; tandem repeat

Query Match          98.9%; Score 91; DB 2; Length 416;
Best Local Similarity 94.7%; Pred. No. 1.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTRGSRGSRGSGS 19
DB 7 SHOESTRGSRGSRGSGS 25

RESULT 2
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C>Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A>Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GAN>
A:Cross-references: GB:J02929
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
```

C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:246-569/Region: filaggrin repeat
 F:570-893/Region: filaggrin repeat
 F:1074-1397/Region: filaggrin repeat
 F:1573-1896/Region: filaggrin repeat

Query Match 98.9%; Score 91; DB 2; Length 2248;
 Best Local Similarity 94.7%; Pred. No. 7.3e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 1055 SHOESTRGRSGRSGSGS 1073

RESULT 3

A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101069; GB:M90967; NID:g190408; PIDN:AAA60177.1; PID:g553621
 A:Note: sequence extracted from NCBI backbone (NCBIP:118773)
 C:Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-q22
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 89.1%; Score 82; DB 2; Length 591;
 Best Local Similarity 89.5%; Pred. No. 5.1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 449 SHOESTRGRSGRSGSGS 467

RESULT 4

W2WL47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D35324
 R:Kiyono, T.; Adachi, A.; Iehibashi, M.
 Virology 177, 401-405, 1990
 A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
 A:Reference number: A53324; MUID:90281611; PMID:2162112
 A:Accession: D35324
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M2305; NID:g333062; PIDN:AAA46979.1; PID:g333067
 C:Superfamily: Papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.22;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 |||||

DB 342 SRENTGRGRGGRGGRGS 360

RESULT 5

T33997
 hypothetical protein W03G1.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33997
 R:Pauley, A.; Scheet, P.; Harper, M.
 submitted to the EMBL Data Library, February 1999
 A:Description: The sequence of C. elegans cosmid W03G1.
 A:Reference number: 221454
 A:Accession: T33997
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-471 <PAU>
 A:Cross-references: EMBL:AF125964; PIDN:AA014753.1; GSPDB:GN00022; CESP:W03G1.5
 A:Experimental source: strain Bristol N2; clone W03G1
 C:Genetics:
 A:Gene: CESP:W03G1.5
 A:Map position: 4

Query Match 58.7%; Score 54; DB 2; Length 471;
 Best Local Similarity 57.9%; Pred. No. 0.84;
 Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 167 SRSFSGRGRGGRSGSGS 185

RESULT 6

T16234
 hypothetical protein F32A5.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16234
 R:Pauley, A.
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of C. elegans cosmid F32A5.
 A:Reference number: 218482
 A:Accession: T16234
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-123 <PAU>
 A:Cross-references: EMBL:U20864; NID:g669026; PIDN:AA046661.1; CESP:F32A5.7
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:F32A5.7
 A:Introns: 2/3; 49/3; 75/3

Query Match 54.3%; Score 50; DB 2; Length 123;
 Best Local Similarity 56.2%; Pred. No. 0.98;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTRGRSGRSGRSG 18
 |||||
 DB 93 RQSRGRSGRSGRSG 108

RESULT 7

T02852
 probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)
 C:Species: Leishmania major
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C:Accession: T02852; H81462
 R:Myler, P.J.
 submitted to the EMBL Data Library, May 1998
 A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
 A:Reference number: Z14740
 A:Accession: T02852
 A:Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-800 <MTL>
A/Cross-references: EMBL:AB001274; NID:g3264850; PID:g2266920
R/Wyler, P.J.; Audleman, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A/Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A/Reference number: AB1455; MIMD:9918987; PMID:10077609
A/Accession: H81462
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-800 <PYL>
A/Cross-references: GB:AE001274; NID:g3264850; PIDN:AA24675.1; PID:g2266920; GSPDB:GN00
A/Experimental source: strain MHOM/IL/81/Friedlin
C/Genetics:
A/Gene: L1439.4
A/Map position: 1

Query Match 54.3%; Score 50; DB 2; Length 800;
Best Local Similarity 55.6%; Pred. No. 5.7;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTRGSKXGRSGS 19
DB 429 HRDVGRLSTTASGRSGS 446

RESULT 8
TS1049
related to nuclear phosphoprotein [imported] - Neurospora crassa
N/Alternate names: protein B12F1.10
C/Species: Neurospora crassa
C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
R/Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A/Accession: T51049
A/Reference number: Z25286
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-822 <SCH>
A/Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12F1.10
A/Experimental source: BAC clone B12F1; strain CR74A
C/Genetics:
A/Gene: NCSP:B12F1.10
A/Map position: 6
A/Introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
Best Local Similarity 52.8%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGSKXGRSGS 19
DB 414 RESASGRTRGRGRGRT 430

RESULT 9
T18918
hypothetical protein C04G2.8 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T18918
R/Hembry, C.
submitted to the EMBL Data Library, April 1996
A/Reference number: Z19045
A/Accession: T18918
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-197 <MTL>
A/Cross-references: EMBL:Z70718; PIDN:CAA94670.1; GSPDB:GN00022; CESP:C04G2.8
A/Experimental source: clone C04G2
C/Genetics:
A/Gene: CESP:C04G2.8

A/Map position: 4
A/Introns: 70/2

Query Match 50.0%; Score 46; DB 2; Length 197;
Best Local Similarity 52.6%; Pred. No. 6.3;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSKXGRSGS 19
DB 104 SRRSSRGSVYGRTRGRS 122

RESULT 10
C87801
protein C10G11.9 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: C87801
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2015, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MIMD:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_eleg
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: C87801
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-203 <STO>
A/Cross-references: GS:chr_1; PIDN:ABA42250.1; PID:g1825626; GSPDB:GN00019; CESP:C10G11.5
C/Genetics:
A/Gene: C10G11.9
A/Map position: 1

Query Match 50.0%; Score 46; DB 2; Length 203;
Best Local Similarity 52.6%; Pred. No. 6.5;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSKXGRSGS 19
DB 110 SRRSSRGSVYGRTRGRS 128

RESULT 11
T21220
hypothetical protein F21H7.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C/Accession: T21220
R/Gardner, A.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19393
A/Accession: T21220
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-306 <MTL>
A/Cross-references: EMBL:Z93379; PIDN:CAB07568.1; GSPDB:GN00023; CESP:F21H7.5
A/Experimental source: clone F21H7
C/Genetics:
A/Gene: CESP:F21H7.5
A/Map position: 5
A/Introns: 28/3; 146/3; 232/3
C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 50.0%; Score 46; DB 2; Length 306;
Best Local Similarity 44.4%; Pred. No. 9.5;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSKXGRSGS 18
DB 48 SHRSKTKNKGSGKSG 65

RESULT 12

T02976
 Probable DNA binding protein PCF2 - rice
 C/Species: Oryza sativa (rice)
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C/Accession: T02976
 R/Kosugi, S.; Ohashi, Y.
 Plant Cell 9, 1607-1619, 1997
 A/Title: PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell
 A/Reference number: 214803; MUID:97480096; PMID:9338963
 A/Accession: T02976
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 15373 <KOS>
 A/Cross-references: EMBL:D87261; NID:G2580439; PIDN:BA23143.1; PID:G2580440
 A/Experimental source: cultivar Nipponbare

Query Match 50.0%; Score 46; DB 2; Length 373;
 Best Local Similarity 53.3%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSG 15
 DB 334 SHEQRRGRKRGNG 348

RESULT 13

hypothetical protein DKFZp761E0323.1 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
 C/Accession: T46259
 R/Bloecher, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A/Reference number: 223032
 A/Accession: T46259
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-399 <AAA>
 A/Cross-references: EMBL:AL137423
 A/Experimental source: adult amygdala; clone DKFZp761E0323
 C/Genetics:
 A/Note: DKFZp761E0323.1

Query Match 50.0%; Score 46; DB 2; Length 399;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSG 18
 DB 336 SNSRQSGWSGSGRSG 353

RESULT 14

probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: G84727
 R/Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nieman, W.C.; White, C.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-769, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A64420; MUID:20083487; PMID:10617197
 A/Accession: G84727
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-836 <STO>
 A/Cross-references: GB:AE002093; NID:G4263718; PIDN:AD15404.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: AC2G32000
 A/Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
 Best Local Similarity 69.2%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 7 RGRSGRSGRSG 19
 DB 812 RGRSGRSGRSG 824

RESULT 15

hypothetical protein F24B22.190 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C/Accession: T47582
 R/Bloecher, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quettier, F.; Salanoubat, M.
 submitted to the Protein Sequence Database, January 2000
 A/Reference number: 223016
 A/Accession: T47582
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1105 <BIO>
 A/Cross-references: EMBL:AL132957
 A/Experimental source: cultivar Columbia; BAC clone F24B22
 C/Genetics:
 A/Map position: 3
 A/Introns: 35/3; 56/2; 294/3; 318/3; 349/3; 376/2; 426/3; 455/1; 485/3; 508/3; 568/3; 633
 A/Note: F24B22.190

Query Match 50.0%; Score 46; DB 2; Length 1105;
 Best Local Similarity 57.9%; Pred. No. 32;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSGRSG 19
 DB 250 SRSRSGRSGRSGRSG 268

Search completed: September 28, 2004, 06:14:58
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 / Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-3

Perfect score: 92
Sequence: 1 SHQESTRGSRXGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	91	98.9	1	FILH_HUMAN
2	58	63.0	2	VE2_HPV47
3	50	54.3	3	LSM4_CABEL
4	45	48.9	4	DEXT_STRDO
5	43	46.7	5	EBP2_CABEL
6	43	46.7	6	FUS_BOVIN
7	43	46.7	7	FUS_MOUSE
8	43	46.7	8	FUS_HUMAN
9	43	46.7	9	YOO5_CABEL
10	43	46.7	10	ICP4_HSVMG
11	43	46.7	11	NKCR_MOUSE
12	42	45.7	12	YDPC_SCHPO
13	42	45.7	13	FRZ2_DROME
14	42	45.7	14	FRQ_CRESE
15	42	45.7	15	SK12_YEAST
16	42	45.7	16	COAC_YEAST
17	41	44.6	17	CMC1_EIMAC
18	41	44.6	18	RIX21_DROME
19	41	44.6	19	2ABA_YEAST
20	41	44.6	20	MR11_CABEL
21	41	44.6	21	SES_RAT
22	41	44.6	22	SON_MOUSE
23	41	44.6	23	SON_HUMAN
24	40.5	44.0	24	RNPH_MYCTU
25	40.5	44.0	25	AN3_XENLA
26	40.5	44.0	26	FBR1_MOUSE
27	40	43.5	27	AP50_DICDI
28	40	43.5	28	NKX2_MOUSE
29	39.5	42.9	29	SPS_SOLTU
30	39	42.4	30	L2MU_ADB40
31	39	42.4	31	DH1_MAIZE
32	39	42.4	32	GRP1_ARTSA
33	39	42.4	33	artemia sal

ALIGNMENTS

RESULT 1	ID	FILE_HUMAN	STANDARD:	PRT:	416 AA.
AC	P20930	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Flaggrin precursor (Fragment).				
GN	FLAG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=89296901; PubMed=2740331;				
RA	Wicklinley-Grant L.V., Idler W.W., Bernstein I.A., Parry D.A.D.,				
RA	Canalizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinhert P.M.;				
RT	"Characterization of a cDNA clone encoding human flaggrin and				
RT	localization of the gene to chromosome region 1q21."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).				
RN	[2]				
RP	CITRULLINATION.				
RX	MEDLINE=96574389; PubMed=8780679;				
RA	Senshu T., Kan S., Ogawa H., Manabe M., Asaga H.;				
RT	"Preferential deimination of keratin K1 and flaggrin during the				
RT	terminal differentiation of human epidermis."				
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).				
CC	-I- FUNCTION: Aggregates keratin intermediate filaments and promotes				
CC	disulfide-bond formation among the intermediate filaments during				
CC	terminal differentiation of mammalian epidermis.				
CC	-I- PTM: Flaggrin is initially synthesized as a large, insoluble,				
CC	highly phosphorylated precursor containing many tandem copies of				
CC	324 AA, which are not separated by "large linker". The precursor				
CC	is deposited as keratohyalin granules. During terminal				
CC	differentiation it is dephosphorylated and proteolytically				
CC	cleaved.				
CC	-I- PTM: Undergoes deimination of some arginine residues				
CC	(citrullination).				
CC	-I- PTM: SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL collaboration				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M24355; AAA52454.1; ..				
DR	PIR; A32947; A32947.				
DR	Genew; HGNC:3748; FLAG.				
DR	MIM; 135940; ..				
DR	GO; GO:0005882; C:intermediate filament; NAS.				
DR	GO; GO:0005198; F:structural molecule activity; NAS.				
DR	GO; GO:0007275; P:development; NAS.				
DR	InterPro; IPR003303; Flaggrin.				

34	39	42.4	315	1	SOL2_YEAST	P37262 saccharomyc
35	39	42.4	325	1	FRZB_BOVIN	Q95117 bos taurus
36	39	42.4	484	1	CLK1_HUMAN	P49759 homo sapien
37	39	42.4	562	1	MYCB_MOUSE	O96c31 mycobacteri
38	39	42.4	638	1	K220_HUMAN	O01546 homo sapien
39	39	42.4	983	1	Y144_HUMAN	Q14157 homo sapien
40	39	42.4	1130	1	SN3B_HUMAN	O75182 homo sapien
41	39	42.4	1527	1	ARHB_RAT	O96667 rattus norv
42	38.5	41.8	135	1	PRH1_DROME	O02427 drosophila
43	38.5	41.8	223	1	PRH1_MOUSE	P97458 mus musculu
44	38.5	41.8	259	1	RNPH_MOUSE	P37939 mycobacteri
45	38.5	41.8	375	1	SR55_DROME	P26666 drosophila

DR Pfam: PF03516; Flaggrin; 2.
 DR PRINTS: P000487; FLAGGRIN.
 KM Phosphorylation; Citrullination; Developmental protein.
 FT NON-TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 98.9%; Score 91; DB 1; Length 416;
 Best Local Similarity 94.7%; Pred. No. 4e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGRSGS 19
 DB 7 SHQSTRGRSGRSGRSGS 25

RESULT 2
 VE2_HP47 STANDARD; PRT; 506 AA.
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 RT type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405 (1990).
 CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2E RESPONSE ELEMENT (5'-ACGNNNNAGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2E'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M32305; AAA46979.1; -.
 DR PIR: D35324; W2WL47.
 DR HSSP: P03122; 2BOP.
 DR InterPro: IPR000427; E2_C.
 DR InterPro: IPR001866; E2_N.
 DR InterPro: IPR009021; Viral_DNA_bd.
 DR Pfam: PF00511; E2_C; 1.
 DR Pfam: PF00508; E2_N; 1.
 DR ProDom: PD000672; E2_C; 1.
 DR ProDom: PD000678; E2_N; 1.
 KW Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF575065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.077;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGRSGS 19

DB 342 SREGNTRGRGRGRGRGAGS 360

RESULT 3
 LSM4_CABE1 STANDARD; PRT; 123 AA.
 AC Q19952;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable U6 snRNA-associated Sm-like protein LSM4.
 GN F32A5.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
 CC snRNA (By similarity).
 CC -1- SUBUNIT: LSM subunits form a dodecamer with a doughnut shape (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U0864; AAC46661.1; -.
 DR PIR: T16234; T16234.
 DR WormPeP: F32A5.7; CE01277.
 DR InterPro: IPR006649; snRNP.
 DR InterPro: IPR001163; snRNP_Sm.
 DR Pfam: PF01423; LSM; 1.
 DR ProDom: PD020287; snRNP; 1.
 DR SMART: SM00651; Sm; 1.
 KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
 KW RNA-binding.
 SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 54.3%; Score 50; DB 1; Length 123;
 Best Local Similarity 56.2%; Pred. No. 0.3;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGRSGRSGRSGS 18
 DB 93 REOSRGRGGGRGGRG 108

RESULT 4
 DEXT_STRDO STANDARD; PRT; 1337 AA.
 ID DEXT_STRDO
 AC P39653;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
 DE glucanohydrolase).
 GN DEX.
 OS Streptococcus downei (Streptococcus sobrinus).
 OG Plaemid PYA902.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1317;

[1] SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.

RP STRAIN=6715 / UAB66;
 RX MEDLINE=94292401; PubMed=8021165;
 RA Wanda S.-Y., Curtiss R. III;
 RT "Purification and characterization of Streptococcus sobrinus
 dectranase produced in recombinant Escherichia coli and sequence
 analysis of the dectranase gene";
 RT J. Bacteriol. 176:3839-3850(1994).
 RL J. FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
 CC PELTICLE-COATED TOOTH SURFACE.
 CC CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
 CC linkages in dextran.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (potential).
 CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
 CC and at 39 degrees Celsius.
 CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M96978; AAA21772.1; -;
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 DR PROSITE; PSS0847; GRAM_POS_ANCHORING; 1.
 KW Hydrolyase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
 KM Signal; Plasmid.
 FT SIGNAL 1 30
 FT CHAIN 1 1308 DEXTRANASE.
 FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
 FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).
 FT MOD RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64.

Query Match 48.9%; Score 45; DB 1; Length 1337;
 Best Local Similarity 47.1%; Pred. No. 24;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRXGRSGRS 17
 Db 1154 ANODSTKSSADSGRS 1170

RESULT 5
 EBP2_CABEL
 ID EBP2_CABEL STANDARD; PRT; 340 AA.
 AC Q09958;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable rRNA processing protein EBP2 homolog.
 GN C18A3.3.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 CC Rhabditidae; Peleodermidae; Caenorhabditis.
 CC NCB1_TaxID=6239;
 RN NCB1 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Hallsworth K.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Waterston R.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Required for the processing of the 27S pre-rRNA (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -1- SIMILARITY: Belongs to the EBP2 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; U28944; AA68370.2; -;
 DR WormRep; C18A3.3; CE28850.
 DR InterPro; IPR008610; EBP2.
 DR Pfam; PF05890; EBP2; 1.
 KW Hypothetical protein; Ribosome biogenesis; Nuclear protein;
 KW Coiled coil.
 KW DOMAIN 206 245 COILED COIL (POTENTIAL).
 SQ SEQUENCE 340 AA; 38083 MW; 13BE25D965A5D55 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 340;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSGRGRSGRS 18
 Db 325 RGRSGRGRSGRS 336

RESULT 6
 FUS_BOVIN
 ID FUS_BOVIN STANDARD; PRT; 512 AA.
 AC Q28609;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE RNA-binding protein FUS (Pisgen protein).
 GN FUS.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 CC Bovidae; Bovinae; Bos.
 CC NCB1_TaxID=9913;
 RN NCB1 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96175600; PubMed=8631501;
 RA Allegro M.C., Allegro M.A.;
 RT "A nuclear protein regulated during the transition from active to
 RT quiescent phenotype in cultured endothelial cells.";
 RL Dev. Biol. 174:288-297(1996).
 RN [2]

RP CARBOHYDRATE-BINDING DOMAIN
 RX MEDLINE=20160719; PubMed=10694442;
 RA Allegro M.C.;
 RT "A C-terminal carbohydrate-binding domain in the endothelial cell
 RT regulatory protein, Pisgen: new function for an EWS family member.";
 RL Exp. Cell Res. 255:270-277(2000).
 CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
 CC promotes ATP-independent annealing of complementary single-
 CC stranded DNAs and D-loop formation in superhelical double-stranded
 CC DNA. May play a role in maintenance of genomic integrity (By
 CC similarity).
 CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Binds SF1 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear; exhibits diffuse staining
 CC throughout (excluding nucleoli), together with a small number of
 CC intensely stained focal points, or granules, and punctate staining
 CC along the nuclear envelope.
 CC -1- DOMAIN: The C-terminal domain binds carbohydrates.

```

CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U26024; AAC13543.1; -.
CC HSRP; P09651; 1HA1.
CC InterPro; IPR000504; RNA_rec_mot.
CC InterPro; IPR001876; Znf_RanGDP.
CC Pfam; PF00076; rrm; 1.
CC Pfam; PF00641; zf-RanBP; 1.
CC SMART; SM00360; RRM; 1.
CC SMART; SM00547; Znf_RBZ; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
CC PROSITE; PS01358; zf_RANBP2_1; 1.
CC PROSITE; PS50199; zf_RANBP2_2; 1.
CC RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
CC Metal-binding.
CC FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.
CC FT DOMAIN 165 253 GLY-RICH.
CC FT DOMAIN 271 357 RNA-BINDING (RRM).
CC FT DOMAIN 357 512 ARG/GLY-RICH.
CC FT ZN_FING 408 439 RANBP2-TYPE.
CC SQ SEQUENCE 512 AA; 52240 MW; 3652339C044F1386 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 512;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSXGRGGRSG 18
Db 228 RGRGGRGGRG 239

RESULT 7
FUS_MOUSE STANDARD; PRT; 518 AA.
AC P56559;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE RNA-binding protein FUS (Fusgen protein).
GN FUS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (PubMed=1295080; PubMed=22829580;
RX MEDLINE=22829580; Zhang M., Zhao X., Allegro M.A., Allegro M.C.,
RA Alappat S.R., Zhang M., Zhao X., Allegro M.A., Allegro M.C.,
RA Burdall C.A.;
RT "Mouse pipen encodes a nuclear protein whose expression is
RT developmentally regulated during craniofacial morphogenesis.";
RT Dev. Dyn. 228:59-71(2003).
CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
CC promotes ATP-independent annealing of complementary single-
CC stranded DNAs and D-loop formation in superhelical double-stranded
CC DNA. May play a role in maintenance of genomic integrity (By
CC similarity).
CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts
CC with Irf3 and Sp1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

```

```

CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF224264; AA170602.1; -.
CC HSRP; P09651; 1HA1.
CC MCD; MGI:135363; Fus.
CC InterPro; IPR000504; RNA_rec_mot.
CC InterPro; IPR001876; Znf_RanGDP.
CC Pfam; PF00076; rrm; 1.
CC Pfam; PF00641; zf-RanBP; 1.
CC SMART; SM00360; RRM; 1.
CC SMART; SM00547; Znf_RBZ; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
CC PROSITE; PS01358; zf_RANBP2_1; 1.
CC PROSITE; PS50199; zf_RANBP2_2; 1.
CC RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
CC Metal-binding.
CC FT DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.
CC FT DOMAIN 168 265 GLY-RICH.
CC FT DOMAIN 278 364 RNA-BINDING (RRM).
CC FT DOMAIN 364 518 ARG/GLY-RICH.
CC FT ZN_FING 415 446 RANBP2-TYPE.
CC SQ SEQUENCE 518 AA; 52673 MW; E06F231BFEED78D6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 518;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSXGRGGRSG 18
Db 235 RGRGGRGGRG 246

RESULT 8
FUS_HUMAN STANDARD; PRT; 526 AA.
AC P35637;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE RNA-binding protein FUS (Oncogene TLS) (Translocated in
DE liposarcoma protein) (POMP75) (75 kDa DNA-pairing protein).
GN FUS OR TLS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93288139; PubMed=8510758;
RA Crozat A., Aman P., Mandahl N., Ron D.;
RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid
RT liposarcoma.";
RT Nature 363:640-644(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=93350637; PubMed=7503811;
RA Rabbits T.H., Foster A., Larson R., Nathan P.;
RT "Fusion of the dominant negative transcription regulator CHOP with a
RT novel gene FUS by translocation t(12;16) in malignant liposarcoma.";
RT Nat. Genet. 4:175-180(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RX MEDLINE=99013873; PubMed=9795213;
RA Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,

```

RA Munakata N., Ohki M.;
 RT "Genomic structure of the human RBP56/hRAF168 and FUS/TLS genes.";
 RL Gene 221:191-198(1998).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Lung; and Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.D., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Munry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Ranev U., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, AND
 RC CHARACTERIZATION.
 RX MEDLINE=20036580; PubMed=10567410;
 RA Bechtold H., Kuroda M., Sok J., Ron D., Lopez B.S., Akhmedov A.T.;
 RT "Human 75-KDa DNA-pairing protein is identical to the pro-oncoprotein
 TLS/FUS and is able to promote D-loop formation.";
 RL J. Biol. Chem. 274:34337-34342(1999).
 RN [6]
 RP SEQUENCE OF 265-276; 317-331 AND 335-357, AND IDENTIFICATION.
 RX MEDLINE=99369251; PubMed=10442642;
 RA Bertrand P., Akhmedov A.T., Delacore F., Durrbach A., Lopez B.S.;
 RT "Human POMP/5 is identified as the pro-oncogene TLF/FUS: both
 POMP/5 and POMP/100 DNA homologous pairing activities are associated
 RT to cell proliferation.";
 RL Oncogene 18:4515-4521(1999).
 RN [7]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE=94243799; PubMed=8187069;
 RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.,
 RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human
 RT myeloid leukemia with t(16;21) chromosomal translocation.";
 RL Cancer Res. 54:2865-2868(1994).
 RN [8]
 RP INTERACTION WITH SP1.
 RX MEDLINE=98325009; PubMed=9660765;
 RA Zhang D., Paley A.J., Childs G.;
 RT "The transcriptional repressor ZFM1 interacts with and modulates the
 RT ability of EMS to activate transcription.";
 RL J. Biol. Chem. 273:18086-18091(1998).
 CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
 CC promotes ATP-independent annealing of complementary single-
 CC stranded DNAs and D-loop formation in superhelical double-stranded
 CC DNA. May play a role in maintenance of genomic integrity.
 CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts
 CC with ILF3 and SP1.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Bvent=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P35637-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P35637-2; Sequence=VSP_005798;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DISEASE: Involved in a form of malignant myeloid liposarcoma
 CC through a chromosomal translocation t(12;16)(q13;p11) that

CC INVOLVES DDIT3 AND FUS.
 CC -1- DISEASE: Involved in a form of acute myeloid leukemia (AML)
 CC through a chromosomal translocation t(16;21)(p11;q22) that
 CC involves FUS and ERG.
 CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL: S62140; AAB27102.1; -;
 CC EMBL: S62138; AAB27103.1; ALT_SEQ.
 CC EMBL: X71427; CA50558.1; ALT_SEQ.
 CC EMBL: X71428; CA50559.1; ALT_SEQ.
 CC EMBL: AF071213; AAC35285.1; -;
 CC EMBL: AF071213; AAC35284.1; -;
 CC EMBL: BC000402; AAH00402.1; -;
 CC EMBL: BC002459; AAH02459.1; -;
 CC PIR: S33799; S33799.
 CC HSRP: P09651; IHA1.
 CC GeneW: HGNC:4010; FUS.
 CC GK: P35637; -;
 CC MIM: 151907; -;
 CC MIM: 151900; -;
 CC GO: GO:0005634; C:nucleus; TAS.
 CC GO: GO:0003723; F:RNA binding; TAS.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC InterPro: IPR001876; Znf_RanBP.
 CC Pfam: PF00076; rtm; 1.
 CC Pfam: PF00641; ZF-RanBP; 1.
 CC SMART: SMO0360; RRM; 1.
 CC SMART: SMO0547; ZNF_RBP2; 1.
 CC PROSITE: PS50102; RRM; 1.
 CC PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC PROSITE: PS01358; ZF_RANBP2_1; 1.
 CC PROSITE: PS01358; ZF_RANBP2_2; 1.
 CC KX Proco-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;
 CC KW Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;
 CC Metal-binding.
 CC FT DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.
 CC FT DOMAIN 166 267 GLY-RICH.
 CC FT DOMAIN 285 371 RNA-BINDING (RRM).
 CC FT DOMAIN 371 526 ARG/GLY-RICH.
 CC FT ZN_FING 422 453 RANBP2-TYPE.
 CC FT SITE 262 267 BREAKPOINT FOR TRANSLOCATION TO FORM
 CC FT VAREPLIC 64 65 FUS/TLS-CHOP ONCOGENE.
 CC FT VAREPLIC 65 65 TG -> S (in isoform short).
 CC FT CONFLICT 338 338 /FTId=VSP_005798.
 CC FT SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;
 CC SQ
 CC Query Match 46.7%; Score 43; DB 1; Length 526;
 CC Best Local Similarity 66.7%; Pred. No. 18;
 CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC QY 7 RGRSYGRSGRSG 18
 CC Db 242 RGRGGRGGRGGRG 253
 CC RESULT 9
 CC Y005_CAEEL
 CC ID Y005_CAEEL STANDARD; PRT; 1232 AA.
 CC AC P34643;
 CC DT 01-FEB-1994 (Rel. 28, Created)
 CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)

DR HSSP, Q27450; 1A33.
 DR MGD; MGI:97346; Nktr.
 DR InterPro; IPR002130; CSA_PPIase.
 DR Pfam; PF00160; pro_isomerase; 1.
 DR PRINTS; PR00153; CSAPRISMASE.
 DR PROSITE; PS00170; CSA_PPIASE_1; 1.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 KM Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
 FT DOMAIN 1 176
 FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
 FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
 FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
 FT DOMAIN 198 273 ARG/SER-RICH.
 FT DOMAIN 468 565 ARG/SER-RICH.
 FT DOMAIN 658 812 ARG/SER-RICH.
 FT DOMAIN 1303 1453 ARG-SER TANDEM REPEAT-RICH.
 SQ SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;

Query Match 46.7%; Score 43; DB 1; Length 1453;
 Best Local Similarity 81.8%; Pred. No. 54;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 RSXGRSGRSGS 19
 Db 682 RSSGSGRSGS 692

RESULT 12
 YDPC_SCHPO STANDARD; PRT; 150 AA.
 ID YDPC_SCHPO
 AC O14015;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C29A4.12c in chromosome I.
 GN SPAC29A4.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajendram M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeert G., Aert R., Robben J., Gymnopoulos B.,
 RA Welfens I., Vansteelt E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesli D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt D., Pohl T.M.,
 RA Beer P., Zimmermann W., Medler H., Wambut R., Pirnalle B.,
 RA Goffeau A., Cadieu R., Driano S., Gloux S., Jellare V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,
 RA Spinkovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL; Z97210; CAB10139.1; -.
 DR FIR; T38477; T38477.
 DR GeneDB; SPCmbe; SPAC29A4.12c; -.
 KW Hypothetical protein.
 SQ SEQUENCE 150 AA; 16781 MW; 3CFPCDE223106361 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 150;
 Best Local Similarity 56.2%; Pred. No. 6.8;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGRSXGRSGR 16
 Db 76 SRSESTRRSGHTR 91

RESULT 13
 FRZ2_DROME STANDARD; PRT; 694 AA.
 ID FRZ2_DROME
 AC Q9VXX3; Q94916; Q9VXX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frizzled protein 2 precursor (Frizzled-2) (dfz2).
 GN FZ2 OR CG9739/CG14083.
 GN FZ2 OR CG9739/CG14083.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=9653971; PubMed=8717036;
 RA Bhant P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.,
 RT "A new member of the frizzled family from Drosophila functions as a
 RT Wingless receptor".
 RL Nature 382:225-230(1996).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abayaratne A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu R., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Docson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.T., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Grotok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Hatvey D.A., Heiman T.U., Hernandez J.R., Houck J.,
 RA Hosini D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Iasko F., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod K.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
CC are coupled to the beta-catenin canonical signaling pathway, which
CC leads to the activation of dishevelled proteins, inhibition of
CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
CC of Wnt target genes. A second signaling pathway involving PKC and
CC calcium fluxes has been seen for some family members, but it is
CC not yet clear if it represents a distinct pathway or if it can be
CC integrated in the canonical pathway, as PKC seems to be required
CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
CC to involve interactions with G-proteins. Required to coordinate
CC the cytoskeletons of epidermal cells to produce a parallel array
CC of cuticular hairs and bristles.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
CC between 15 and 70 per cent of egg length, including the
CC invaginating cells of the ventral furrow. Stripe pattern is
CC emerging by early stage 8. From stage 9 and continuing throughout
CC embryogenesis, expression is seen in the developing CNS. At stage
CC 10, expressed in 15 stripes in the presumptive head and trunk
CC regions, in the posterior midgut primordium, in a subset of cells
CC of anterior midgut invagination and in the procephalic lobe. At
CC stage 12, expression declines in epidermis and increases in the
CC midgut and visceral mesoderm. At stage 17, only expressed in the
CC CNS, hindgut and dorsal vessel.
CC -1- DOMAIN: Lys-Thr-X-X-TRP motif is involved in the activation of
CC the Wnt/beta-catenin signaling pathway (By similarity).
CC -1- DOMAIN: The FZ domain is involved in binding with Wnt ligands.
CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
CC family.
CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U65589; AAC47273.1; -;
CC EMBL: AE003518; AAF49185.2; -;
CC PIR: S71786; S71786;
CC FLYBase: FBgn0016797; fz2.
CC GO: GO:0016021; C: integral to membrane, NAS.
CC GO: GO:0017147; F: Wnt-protein binding, IDA.
CC GO: GO:0006928; P: cell motility, IMP.
CC GO: GO:0007163; P: establishment and/or maintenance of cell po. .; ISS.
CC GO: GO:0008585; P: female gonad development, IMP.
CC GO: GO:0016055; P: Wnt receptor signaling pathway, IDA.
CC InterPro: IPR000539; Frizzled.
CC InterPro: IPR000024; Fz domain.
CC InterPro: IPR000832; GPCR secretin.
CC Pfam: PF01534; Frizzled, 1.
CC Pfam: PF01392; Fz, 1.
CC PRINTS: PR00489; FRIZZLED.
CC SMART: SM00063; FRT, 1.
CC PROSITE: PS50038; FZ, 1.
CC PROSITE: PS50261; G_PROTEIN_RECEP_F2_4, 1.
CC Multigene family: Receptor; G-protein coupled receptor; Transmembrane;
CC Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.

FT SIGNAL 1 22
FT CHAIN 23 694
FT DOMAIN 23 315
FT TRANSMEM 316 336
FT DOMAIN 337 352
FT TRANSMEM 353 373
FT DOMAIN 374 397
FT TRANSMEM 398 418
FT DOMAIN 419 439
FT TRANSMEM 440 460
FT DOMAIN 461 482
FT TRANSMEM 483 503
FT DOMAIN 504 534
FT TRANSMEM 535 555
FT DOMAIN 556 584
FT TRANSMEM 585 605
FT DOMAIN 606 694
FT DOMAIN 187 225
FT DOMAIN 59 180
FT SITE 608 613
FT SITE 682 694
FT CARBOHYD 78 78
FT CARBOHYD 288 288
FT CONFLICT 55 55
FT CONFLICT 417 417
SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAB096 CRC64;
Query Match 45.7%; Score 42; DB 1; Length 694;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
DB 1 SHOESTRGRSGSGSGSGS 19
182 SYTEAGSGSGSGSGSGSGS 200
RESULT 14
PRO_CRESP ID PRO_CRESP STANDARD; PRT; 1015 AA.
AC 000586;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Frequency clock protein.
GN PRO.
OS Creopus spinulosus (Chromocorea spinulosa).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=110619;
RX SEQUENCE FROM N.A.
RC STRAIN=FGSC 4896;
RX MEDLINE=97188515; PubMed=9037100;
RA Lewis M.T., Morgan L.W., Feldman J.F.;
RT "Analysis of frequency (frq) clock gene homologs: evidence for a
RT helix-turn-helix transcription factor.";
RL Mol. Gen. Genet. 253:401-414(1997).
CC -1- FUNCTION: Circadian clock component involved in the generation of
CC biological rhythms, in particular in rhythm stability, period
CC length, and temperature compensation. Behaves as a negative
CC element in circadian transcriptional loop (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U25850; AAA88072.1; -;
CC PIR: T42013; T42013.

```

KW Biological rhythms; Transcription regulation; Nuclear protein.
FT DOMAIN 240 245 POLY-SER.
FT DOMAIN 356 368 POLY-GIN.
FT DOMAIN 443 451 POLY-PRO.
FT DOMAIN 584 588 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 892 913 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 1015 AA; 110972 MW; EA49E732ED741481 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 53;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
Db NHRKQKTGHSTGDSGSGN 601

RESULT 15
SKI2_YEAST STANDARD; PRT; 1287 AA.
AC P35207; Q06047;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Antiviral protein SKI2.
GN SKI2 OR YLR198C OR L8084.17.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93309467; PubMed=8321235;
RA Widner W.R., Wickner R.B.;
RT "Evidence that the Ski antiviral system of Saccharomyces cerevisiae
acts by blocking expression of viral mRNA.";
RL Mol. Cell. Biol. 13:4331-4341(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RC MEDLINE=97313267; PubMed=9169871;
RA Uomston M., Hillier L., Riles L., Albertmann K., Andre B., Ansoorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Ertian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Nuer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrastarazu L.A., Vandenbol M., Verhaeselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambut R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN [3]
RP SEQUENCE OF 1-162 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95116323; PubMed=7816623;
RA Lygerou Z., Conesa C., Leesele P., Swanson R.N., Ruet A., Carlson M.,
RA Senterac A., Seraphin B.;
RT "The yeast BDF1 gene encodes a transcription factor involved in the
expression of a broad class of genes including snRNAs.";
RL Nucleic Acids Res. 22:5332-5340(1994).
CC BLOCKING TRANSLATION OF VIRAL MRNAs, PERHAPS RECOGNIZING THE
ABSENCE OF CAP OR POLY(A). ESSENTIAL FOR CELL GROWTH ONLY IN THE
PRESENCE OF M1 REPLICON. SEEMS TO BE A HELICASE.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: Belongs to the helicase family. SKI2 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13469; AAA35049.1; -
DR EMBL: U19729; AAB82356.1; -
DR EMBL: Z18944; CAA79378.1; -
DR PIR: S55954; S55954.
DR Geronline; 142461; -
DR SGD: S0004390; SKI2.
DR GO: GO:0005737; Cytoplasm; IDA.
DR GO: GO:0006402; PiRNA catabolism; IMP.
DR GO: GO:0006445; Pi:regulation of translation; IGI.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Antiviral; Transcription regulation; RNA-binding; DNA-binding; Helicase;
KW ATP-binding.
FT NP BIND 351 358 ATP (POTENTIAL).
FT SITE 444 447 DEVH BOX.
FT DOMAIN 556 577 RNA-BINDING RG3-BOX (BY SIMILARITY).
FT DOMAIN 555 597 ARG/GLY-RICH.
FT CONFLICT 326 326 W -> C (IN REF. 1).
FT CONFLICT 759 760 QM -> L (IN REF. 1).
SQ SEQUENCE 1287 AA; 146058 MW; 7CCD36CFC0DF8C32 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 1287;
Best Local Similarity 47.4%; Pred. No. 68;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
Db 550 SKTDNGRGSTRGGRSGS 568

```

Search completed: September 28, 2004, 06:05:47
 Job time : 8.22396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-3

Perfect score: 92
Sequence: 1 SHQESTRGSRSGSRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_rhbc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	798	4 Q9HAU3	Q9HAU3 homo sapien
2	91	98.9	1084	4 Q01212	Q01212 homo sapien
3	90	97.8	797	4 Q16824	Q16824 homo sapien
4	90	97.8	990	4 Q15206	Q15206 homo sapien
5	90	97.8	1218	4 Q05331	Q05331 homo sapien
6	82	89.1	591	4 Q01720	Q01720 homo sapien
7	82	89.1	687	4 Q9HAU2	Q9HAU2 homo sapien
8	77	83.7	465	4 Q03838	Q03838 homo sapien
9	76	82.6	322	4 Q75370	Q75370 homo sapien
10	54	58.7	471	5 Q9UAY0	Q9UAY0 caenorhabdi
11	50	54.3	800	5 Q15845	Q15845 leishmania
12	50	54.3	1284	5 Q9V9Y3	Q9V9Y3 drosophila
13	48	52.2	822	3 Q9P312	Q9P312 neurospora
14	48	52.2	1229	3 Q86ZHI	Q86ZHI neurospora
15	47	51.1	52	16 Q98BE2	Q98BE2 rhizobium 1
16	47	51.1	349	16 Q7UL67	Q7UL67 rhodospirillum

17	47	51.1	436	16 Q82C67	Q82C67 streptomyces
18	46.5	50.5	813	12 Q91IN8	Q91IN8 hepatitis b
19	46	50.0	197	5 Q17626	Q17626 caenorhabdi
20	46	50.0	203	5 P91026	P91026 caenorhabdi
21	46	50.0	306	5 Q45386	Q45386 caenorhabdi
22	46	50.0	373	10 Q23876	Q23876 oryza sativ
23	46	50.0	399	4 Q9NTA9	Q9NTA9 homo sapien
24	46	50.0	467	4 Q9BWI8	Q9BWI8 homo sapien
25	46	50.0	737	4 Q9BO39	Q9BO39 homo sapien
26	46	50.0	836	10 Q9SKZ9	Q9SKZ9 arabidopsis
27	46	50.0	1007	10 Q8VYR8	Q8VYR8 arabidopsis
28	46	50.0	1105	10 Q9M383	Q9M383 arabidopsis
29	45.5	49.5	573	5 Q81GR2	Q81GR2 drosophila
30	45.5	49.5	1910	5 Q9V768	Q9V768 drosophila
31	45	48.9	243	10 Q9AX45	Q9AX45 oryza sativ
32	45	48.9	374	5 Q76561	Q76561 caenorhabdi
33	45	48.9	476	12 Q80890	Q80890 herpesvirus
34	44.5	48.4	319	10 Q9A4J0	Q9A4J0 arabidopsis
35	44	47.8	89	16 Q9A9V2	Q9A9V2 caulobacter
36	44	47.8	229	11 Q9JW93	Q9JW93 mus musculus
37	44	47.8	259	4 Q8VYA2	Q8VYA2 mus musculus
38	44	47.8	282	3 Q871R2	Q871R2 mus musculus
39	44	47.8	333	10 Q9SE88	Q9SE88 brassica o1
40	44	47.8	344	4 Q8N444	Q8N444 homo sapien
41	44	47.8	426	4 Q8V776	Q8V776 homo sapien
42	44	47.8	588	5 Q819N5	Q819N5 schistosoma
43	44	47.8	669	11 Q8BRU5	Q8BRU5 mus musculus
44	44	47.8	734	11 Q9NM09	Q9NM09 mus musculus
45	44	47.8	867	5 Q18218	Q18218 caenorhabdi

ALIGNMENTS

RESULT 1
Q9HAU3 PRELIMINARY: PRT: 798 AA.
AC Q9HAU3
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJL4N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.
RL Submitted (OCF-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; ..
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGRIN.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
Query Match 98.9%; Score 91; DB 4; Length 798;
Best Local Similarity 94.7%; Pred. No. 9.5e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHQESTRGSRSGSRSGS 19
DB 428 SHQESTRGSRSGSRSGS 446
RESULT 2
Q01212 PRELIMINARY: PRT: 1084 AA.
AC Q01212; Q03840; ..
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60501; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED
DR GO; GO:0005882; C:Intermediate filament; NAS.
DR GO; GO:0005198; F:Structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
DR NON TER
FT
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B9D5A362D CRC64;

Query Match 98.9%; Score 91; DB 4; Length 1084;
Best Local Similarity 94.7%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQESTRGRSGRSGSGS 19
Db 64 SHQESTRGRSGRSGSGS 82

RESULT 3
Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60501; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED
DR GO; GO:0005882; C:Intermediate filament; NAS.
DR GO; GO:0005198; F:Structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
DR NON TER
FT
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDAB6B CRC64;

Query Match 97.8%; Score 90; DB 4; Length 797;
Best Local Similarity 94.7%; Pred. No. 1.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQESTRGRSGRSGSGS 19
Db 427 SHQESTRGRSGRSGSGS 445

```

```

RESULT 4
Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60494; AAA63244.1; -.
DR EMBL; M60494; AAA63244.1; JOINED
DR GO; GO:0005198; F:Structural molecule activity; IA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
DR NON TER
FT
SQ SEQUENCE 990 AA; 106453 MW; A8396F10FA91991 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 990;
Best Local Similarity 94.7%; Pred. No. 1.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQESTRGRSGRSGSGS 19
Db 875 SHQESTRGRSGRSGSGS 893

RESULT 5
Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGRIN (PROFILAGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.C., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC -I- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFIDE-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -I- PTM: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF

```

324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 PROTEOLYTICALLY CLEAVED.
 CC POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; M96943; AAA36487.1; -.
 DR PIR; A48118; A48118.
 DR HSSP; P02593; 1CDM.
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0005509; F:calcium ion binding; TAS.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:008151; P:cell growth and/or maintenance; NAS.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 6.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
 KM Polymorphism.
 FT CA BIND 19 32 SITE I (BY SIMILARITY).
 FT CA BIND 62 73 SITE II (BY SIMILARITY).
 FT NON TER 1218 1218
 SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 1218;
 Best Local Similarity 94.7%; Pred. No. 2.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSGS 19
 Db 1097 SHQESTRGSRGSRGSGS 1115

RESULT 6
 ID Q01720 PRELIMINARY; PRT; 591 AA.
 AC Q01720;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Prestlad R.B., Haydock P.V., Fleckman P., Niyonskisi W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 RT organization and identification of an S-100-like calcium binding
 RT domain at the amino terminus.";
 RL J. Biol. Chem. 267:23772-23781(1992).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; L01089; AAA60177.1; -.
 DR EMBL; L01090; AAA60176.1; -.
 DR

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Polymorphism.
 FT PROPEP 1 293 POTENTIAL.
 FT CHAIN 294 467 FILAGGRIN.
 FT PROPEP 468 474 POTENTIAL.
 FT CHAIN 475 >591 FILAGGRIN.
 FT CA BIND 19 32 SITE I (BY SIMILARITY).
 FT CA BIND 62 73 SITE II (BY SIMILARITY).
 FT NON TER 591 591
 SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 89.1%; Score 82; DB 4; Length 591;
 Best Local Similarity 89.5%; Pred. No. 2.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSGS 19
 Db 449 SHQESTRGSRGSRGSGS 467

RESULT 7
 ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
 AC Q9H4U2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laid G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL36504; CAC13172.1; -.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON TER 687 687
 SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 89.1%; Score 82; DB 4; Length 687;
 Best Local Similarity 89.5%; Pred. No. 2.7e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGS 19
 DB 449 SHOESTRGRSGRSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003838;
 DT 01-NOV-1996 (TREMBLrel. 01. Created)
 DT 01-NOV-1996 (TREMBLrel. 01. Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25. Last annotation update)
 DE FILAGRIN (PROFILAGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN (1)

SEQUENCE FROM N.A.

RP TISSUE=PLACENTA;

RC MEDLINE=91064347; PubMed=2248957;

RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin

gene.",

RL Biochemistry 29:9432-9440(1990).

RN (2)

REVIEWS.

RP MEDLINE=91255199; PubMed=2043621;

RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin

gene.",

RL Biochemistry 30:5814-5814(1991).

-1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND

PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

-1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND

WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

REPEATS.

-1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,

INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN

GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

PROTEOLYTICALLY CLEAVED.

EMBL: M60499; AAA63246.1; -

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam: PF03516; Filaggrin; 3.

DR PRINTS: PR00487; FILAGRIN.

FT NON_TER 1 1

FT 465 465

SEQUENCE 465 AA; 50280 MW; C683744C5E134097 CRC64;

Query Match 83.7%; Score 77; DB 4; Length 465;

Best Local Similarity 84.2%; Pred. No. 0.00012;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGS 19

DB 227 SHOESTRGRSGRSGS 245

RESULT 9

075370 PRELIMINARY; PRT; 322 AA.

AC 075370;

DT 01-NOV-1998 (TREMBLrel. 08. Created)

DT 01-NOV-1998 (TREMBLrel. 08. Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24. Last annotation update)

DE Epidermal filaggrin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN (1)

SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;

RA Girbal-Neuhausser E., Durieux J.-J., Arnaud M., Dalpon P., Sebbag M.,

RA Vincent C., Simon M., Sensu T., Masson-Bessiere C.,

RA Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated

antifilaggrin autoantibodies are posttranslationally generated on

RT various sites of (pro)filaggrin by deamination of arginine residues.;"

RL J. Immunol. 162:585-594(1999).

DR EMBL: AF043380; AAC23559.1; -

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam: PF03516; Filaggrin; 2.

DR PRINTS: PR00487; FILAGRIN.

FT NON_TER 1 1

FT 322 322

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match 82.6%; Score 76; DB 4; Length 322;

Best Local Similarity 83.3%; Pred. No. 0.00012;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGS 18

DB 305 SHOESTRGRSGRSGS 322

RESULT 10

09UAY0 PRELIMINARY; PRT; 471 AA.

AC 09UAY0;

DT 01-MAY-2000 (TREMBLrel. 13. Created)

DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24. Last annotation update)

DE W03G1.5 protein.

GN W03G1.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Felodermidae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN (1)

SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Burke M.,

RA Bonfield U., Burton J., Connell M., Copey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,

RA Smailon N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg U., Thomas K., Vaudin K., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat U., Wohlman P.,

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans.;"

RL Nature 368:32-38(1994).

RN (2)

SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RA Pauley A., Scheet P., Harper M.;

RT "The sequence of C. elegans cosmid W03G1.;"

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

[3]

SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF125964; AAD14753.1; -

DR PIR: T33997; T33997.

DR WormPep: W03G1.5; CE17283.


```

SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A985B CRC64;
Query Match 58.7%; Score 54; DB 5; Length 471;
Best Local Similarity 57.9%; Pred. No. 0.92;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSGS 19
   : ||| ||| |||
Db 167 SSRSPRGRGRGRGGRGSGS 185

RESULT 11
ID 015845 PRELIMINARY; PRT; 800 AA.
AC 015845;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE L1439.4.
GN Leishmania major.
OS Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=9178987; Pubmed=10077609;
RA Wyler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
RA Magnus C., Rickel E., Stark E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AB001274; AAC24675.1; -.
DR PIR; T02852; T02852.
SQ SEQUENCE 800 AA; 83782 MW; DA57AE627933CBAC CRC64;

QY Query Match 54.3%; Score 50; DB 5; Length 800;
Best Local Similarity 55.6%; Pred. No. 7.8;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 429 HRDVGRLSTTASGRSGS 446

QY 2 HOESTRGSRGSRGSGS 19
   : ||| ||| |||
ID 09V9Y3 PRELIMINARY; PRT; 1284 AA.
AC 09V9Y3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CG11339-PA.
GN CG11339.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richardson S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang L.X.,
RA Brendon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gaber G.L.,
RA Abilil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Balilew R.M., Baer P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

```

```

RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck G., Gong F., Gorrell U.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mody B., Murphy B., Murphy L., Muzny D.W., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazon D.M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN 12
RP SEQUENCE FROM N.A.
RA Miska S., Crosby M.A., Matthews B.B., Bayraktaroglu I., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celnik S.E.,
RA Ciamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Krommiller B., Marshall B., Millburn G.F., Richter J.,
RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnick F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AE003777; AAF57145.2; -.
DR FlyBase; FBgn003841; CG11339.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_41; 1.
DR PRINTS; PR00395; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00577; FERM_3; 1.
SQ SEQUENCE 1284 AA; 138966 MW; EDF4119838AA6447 CRC64;

QY Query Match 54.3%; Score 50; DB 5; Length 1284;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 545 SENESELSRGSRGSRGSG 562

RESULT 13
ID 09P312 PRELIMINARY; PRT; 822 AA.
AC 09P312;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

```

```

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12P1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algn V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -.
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR005120; Smg_4_UFP3.
DR Pfam; PF000076; rtm; 1.
DR Pfam; PF03467; Smg4_UFP3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 52.2%; Score 48; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QESTRGSRGSRGSGS 19
Db 414 RESAGRRGRGRGCGT 430

RESULT 14
Q86ZHI PRELIMINARY; PRT; 1229 AA.
AC Q86ZHI;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Related to cell division cycle 2-related protein kinase 7.
GN 7P4.060.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algn V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX294020; CAD70910.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0000910; P:cytokinesis; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKC; 1.

```

```

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Cell division; Kinase.
SQ SEQUENCE 1229 AA; 136776 MW; 43BE612656FF73E4 CRC64;

Query Match 52.2%; Score 48; DB 3; Length 1229;
Best Local Similarity 57.9%; Pred. No. 27;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SHQESTRGSRGSRGSGS 19
Db 242 SHHERRSRDRKRSRSGRSRS 260

RESULT 15
Q98BE2 PRELIMINARY; PRT; 52 AA.
ID Q98BE2;
AC Q98BE2;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein ms15615.
GN MS15615.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matenabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003007; BAB52030.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 52 AA; 5836 MW; 505124C0543255FA CRC64;

Query Match 51.1%; Score 47; DB 16; Length 52;
Best Local Similarity 64.3%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 TRGSRGSRGSGS 19
Db 39 TRAGSAGRAGTSGS 52

```

Search completed: September 28, 2004, 06:12:41
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHQESTRGSRXGSRGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_296Jan04:*
1: geneseqp1960s:*
2: geneseqp1950s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	2	AAW61507
2	90	97.8	19	2	AAW61514
3	90	97.8	1467	5	ABB97605
4	86	93.5	19	2	AAW61517
5	85	92.4	19	2	AAW61516
6	84	91.3	19	2	AAW61506
7	84	91.3	19	2	AAW61508
8	84	91.3	19	2	AAW61515
9	84	91.3	19	2	AAW61509
10	84	91.3	19	2	AAW61511
11	84	91.3	19	2	AAW61505
12	78	84.8	19	2	AAW61512
13	78	84.8	19	2	AAW61513
14	78	84.8	19	2	AAW61510
15	75	81.5	21	2	AAW61520
16	75	81.5	22	4	AAE07235
17	66	71.7	330	2	AAV22956
18	66	71.7	330	2	AAV22955
19	66	71.7	330	2	AAV22957
20	62	67.4	330	2	AAV22954
21	57	62.0	477	6	ABO07142
22	54	58.7	1711	4	AAW79819
23	54	58.7	1951	4	AAW79835
24	50	54.3	1199	4	ABBS5274
25	47	51.1	184	3	AAW57041

ALIGNMENTS

26	46	50.0	41	4	AAW16780	AAW16780 Peptide #
27	46	50.0	41	4	ABB35764	ABB35764 Peptide #
28	46	50.0	41	4	AAW29265	AAW29265 Peptide #
29	46	50.0	41	4	ABB30600	ABB30600 Peptide #
30	46	50.0	41	4	ABB21188	ABB21188 Protein #
31	46	50.0	41	4	AAW68954	AAW68954 Human bon
32	46	50.0	41	4	AAW56574	AAW56574 Human bra
33	46	50.0	41	4	ABG50617	ABG50617 Human liv
34	46	50.0	41	4	AAW04497	AAW04497 Peptide #
35	46	50.0	41	5	ABG38540	ABG38540 Human pep
36	46	50.0	443	4	AAW40069	AAW40069 Human pol
37	46	50.0	574	4	AAW41855	AAW41855 Human pol
38	46	50.0	752	4	AAU23535	AAU23535 Novel hum
39	45.5	49.5	1945	4	ABB64947	ABB64947 Drosophill
40	44	47.8	229	7	ADBS9825	ADBS9825 Rat Prote
41	44	47.8	229	7	ADBS9833	ADBS9833 Rat Prote
42	44	47.8	229	7	ADBS9837	ADBS9837 Rat Prote
43	44	47.8	229	7	ADBS9829	ADBS9829 Rat Prote
44	44	47.8	1038	7	ADC03412	ADC03412 Rice flow
45	44	47.8	1308	6	ABR57072	ABR57072 Human tau

RESULT 1
AAW61507
ID AAW61507 standard; peptide; 19 AA.
XX
AC AAW61507;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
OS
OS
FH Key location/Qualifiers
FT Modified-site 11 /note="Citruilline"
FT
XX
XX
XX W09822503-A2.
XX
XX
XX 26-MAY-1998.
XX
XX 14-NOV-1997; 97MO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats OWH, Hoest RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX
XX Disclosure; Page 6; 199p; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGSGS 19
 DB 1 SHOESTRGRSKXGRSGSGS 19

RESULT 2
 AAW61514
 ID AAW61514 standard; peptide; 19 AA.

AC AAW61514;

DT 26-OCT-1998 (first entry)

DE Peptide c.f. based on cDNA of a profilaggrin repeat.

XX Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI, 1998-396613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61500 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.6e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGSGS 19
 DB 1 SHOESTRGRSKXGRSGSGS 19

RESULT 3

ABB97605
 ID ABB97605 standard; protein; 1467 AA.
 AC ABB97605;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

XX Human; anti-anaemic; vulnary; anti-inflammatory; immunomodulator;
 KW anti-fertility; cerebroprotective; cytoprotective; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wenman T, Drmanac RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABN32791.

PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 873; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate actin or inhibit e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

SQ Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.7e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGSGS 19
 DB 1097 SHOESTRGRSKXGRSGSGS 1115

RESULT 4
 AAW61517
 ID AAW61517 standard; peptide; 19 AA.

AC AAW61517;

DT 26-OCT-1998 (first entry)

DE Peptide c.f. based on cDNA of a profilaggrin repeat.

XX Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

```

OS Synthetic.
OS Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 93.5%; Score 86; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.1e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHOESTGRSGRSGRSGS 19
Db 1 SHOESTGRSGRSGRSGS 19
RESULT 5
AAW61516
ID AAW61516 standard; peptide; 19 AA.
XX
XX AAW61516;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfe, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;

```

```

XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 92.4%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTGRSGRSGRSGS 19
Db 1 SHOESTGRSGRSGRSGS 19
RESULT 6
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX
XX AAW61506;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 9
XX FT /note= "Citruiline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 |||||
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 7
 AAW61508
 ID AAW61508 standard; peptide; 19 AA.

XX AAW61508;

DT 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "Citrulline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI, 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 |||||
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 8
 AAW61515
 ID AAW61515 standard; peptide; 19 AA.

XX AAW61515;

DT 26-OCT-1998 (first entry)

XX Peptide cfa, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI, 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 |||||
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 9
 AAW61509
 ID AAW61509 standard; peptide; 19 AA.

XX AAW61509;

DT 26-OCT-1998 (first entry)

XX

DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 16
 FT /note= "Citruiline"
 PN WO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHQESTGRSGRSGRSGS 19
 DB 1 SHQESTGRSGRSGRSGS 19

RESULT 10
 AAM61511
 ID AAM61511 standard; peptide; 19 AA.
 XX
 AC AAM61511;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruiline"
 FT Modified-site 11
 FT /note= "Citruiline"
 PN WO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHQESTGRSGRSGRSGS 19
 DB 1 SHQESTGRSGRSGRSGS 19

RESULT 11
 AAM61505
 ID AAM61505 standard; peptide; 19 AA.
 XX
 AC AAM61505;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc1, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT
 PN WO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHQESTXGRSGRSGSGS 19
 AC |||||
 DB 1 SHQESTXGRSGRSGSGS 19

RESULT 12

AAW61512
 ID AAW61512 standard; peptide; 19 AA.

AC AAW61512;

DT 26-OCT-1998 (first entry)

XX Peptide cfc8, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citnulline"

FT Modified-site 13 /note= "Citnulline"

FT Modified-site /note= "Citnulline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHQESTXGRSGRSGSGS 19
 AC |||||
 DB 1 SHQESTXGRSGRSGSGS 19

RESULT 13

AAW61513
 ID AAW61513 standard; peptide; 19 AA.

AC AAW61513;

DT 26-OCT-1998 (first entry)

XX Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citnulline"

FT Modified-site 16 /note= "Citnulline"

FT Modified-site /note= "Citnulline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTGRSGRSGSGS 19
 1 SHQESTGRSGRSGSGS 19
 Db 1 SHQESTGRSGRSGSGS 19

RESULT 14
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

AC AAW61510;
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cf66, based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

XX MO9822503-A2.
 XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTGRSGRSGSGS 19
 1 SHQESTGRSGRSGSGS 19
 Db 1 SHQESTGRSGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide XI based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

XX MO9822503-A2.
 XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 6.6e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HOESTGRSGRSGSGS 19
 4 HOESTGRSGRSGSGS 21
 Db 4 HOESTGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:12
 Job time : 55.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHQESTRGSRXGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/ECTS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-3
2	86	97.8	19	9	US-09-308-150-11
3	90	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-4
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-7
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-8
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	57	62.0	477	15	US-10-161-927-62	Sequence 62, Appl
17	50	54.3	133	16	US-10-437-963-136367	Sequence 136367,
18	49	53.3	123	16	US-10-767-701-32436	Sequence 32436, A
19	48	52.2	179	12	US-10-425-114-35345	Sequence 5345, A
20	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
21	47	51.1	436	14	US-10-156-761-13022	Sequence 13022, A
22	47	51.1	511	16	US-10-437-963-188040	Sequence 188040,
23	47	51.1	563	16	US-10-437-963-198539	Sequence 198539,
24	47	51.1	889	16	US-10-437-963-181020	Sequence 181020,
25	46	50.0	41	9	US-09-864-161-36486	Sequence 36486, A
26	46	50.0	164	16	US-10-437-963-112419	Sequence 112419,
27	46	50.0	211	14	US-10-002-631C-268	Sequence 268, App
28	46	50.0	373	16	US-10-437-963-143403	Sequence 143403,
29	46	50.0	461	16	US-10-437-963-185969	Sequence 185969,
30	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
31	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
32	45	48.9	227	16	US-10-437-963-181049	Sequence 181049,
33	45	48.9	243	16	US-10-437-963-103033	Sequence 103033,
34	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
35	44	47.8	78	16	US-10-437-963-203073	Sequence 203073,
36	44	47.8	155	16	US-10-437-963-161956	Sequence 161956,
37	44	47.8	445	12	US-10-424-599-267922	Sequence 267922,
38	44	47.8	612	12	US-10-424-599-230819	Sequence 230819,
39	44	47.8	1071	16	US-10-437-963-118389	Sequence 118389,
40	44	47.8	1270	16	US-10-408-765A-4253	Sequence 2253, Ap
41	44	47.8	1321	16	US-10-271-507-2	Sequence 2, Appl
42	43.5	47.3	19723	15	US-10-084-846A-5	Sequence 5, Appl
43	43	46.7	26	13	US-10-056-407-22	Sequence 22, Appl
44	43	46.7	46	13	US-10-056-407-20	Sequence 20, Appl
45	43	46.7	50	14	US-10-029-386-27692	Sequence 27692, A

ALIGNMENTS

RESULT 1
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michel Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150, PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 SHOESTRGRSGRSGSGS 19
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 2
US-09-308-150-11
; Sequence 11, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-11

Query Match      97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 3
US-09-308-150-14
; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-14

Query Match      93.5%; Score 86; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 9.1e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 4
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-13

Query Match      92.4%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 5
US-09-308-150-1
; Sequence 1, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

```

```

; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 6
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS.
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 7
US-09-308-150-4

```

```

; Sequence 4, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 8
US-09-308-150-5
; Sequence 5, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline

```

US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 1 SHOESTRGRSGRSGSGS 19

RESULT 9

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.7%; Pred. No. 1.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 1 SHOESTRGRSGRSGSGS 19

RESULT 10

US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 1 SHOESTRGRSGRSGSGS 19

RESULT 11

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 1 SHOESTRGRSGRSGSGS 19

RESULT 12

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-8

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGXGRSGSGS 19
Db 1 SHOESTXGRSGXGRSGSGS 19

RESULT 13

US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-9

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGXGRSGSGS 19
Db 1 SHOESTXGRSGXGRSGSGS 19

RESULT 14

US-09-308-150-10
Sequence 10, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
NAME/KEY: DISULFID
LOCATION: (3)..(16)
US-09-308-150-10

Query Match 81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00039;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HOESTGRSGXGRSGSGS 19
Db 4 HOESTXGRSGXGRSGSGS 21

RESULT 15

US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydie
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.0031NUS00 INNS:031
CURRENT APPLICATION NUMBER: US/09/747,029A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 22
LENGTH: 22
TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
; NAME/KEY: MOD_RES
; LOCATION: (10)-(10)
; OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

Query Match 81.5%; Score 75; DB 9; Length 22;
Best Local Similarity 83.3%; Pred. No. 0.00041;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Caps 0;

QY 2 HQESTGRSGRSGRSGS 19
Db 5 HQESTGRSGRSGRSGS 22

Search completed: September 28, 2004, 07:28:50
Job time : 112.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 / Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-3

Perfect score: 92

Sequence: 1 SHOESTRGRSGRSGSGS 19

Scoring table: BLOSUM62

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	53.3	629	US-09-252-991A-22901	Sequence 22901, A
2	47	51.1	406	US-09-252-991A-19857	Sequence 19857, A
3	45	48.9	395	US-09-252-991A-21702	Sequence 21702, A
4	45	48.9	778	US-09-252-991A-18605	Sequence 18605, A
5	44	47.8	369	US-09-252-991A-20245	Sequence 20245, A
6	44	47.8	374	US-09-252-991A-33040	Sequence 33040, A
7	44	47.8	483	US-09-252-991A-19015	Sequence 19015, A
8	44	47.8	518	US-09-252-991A-25967	Sequence 25967, A
9	43	46.7	26	US-09-297-981A-22	Sequence 22, Appl
10	43	46.7	46	US-09-297-981A-20	Sequence 20, Appl
11	43	46.7	159	US-09-252-991A-27107	Sequence 27107, A
12	43	46.7	653	US-09-513-057C-13	Sequence 13, Appl
13	42	45.7	85	US-09-252-991A-29339	Sequence 29339, A
14	42	45.7	105	US-09-252-991A-32394	Sequence 32394, A
15	42	45.7	175	US-09-252-991A-21788	Sequence 21788, A
16	42	45.7	803	US-09-252-991A-23614	Sequence 23614, A
17	42	45.7	2237	US-08-354-973-1	Sequence 1, Appl
18	41.5	45.1	237	US-09-252-991A-17927	Sequence 17927, A
19	41.5	45.1	398	US-09-252-991A-19488	Sequence 19488, A
20	41.5	45.1	415	US-09-328-352-6430	Sequence 6430, Ap
21	41	44.6	120	US-09-252-991A-17025	Sequence 17025, A
22	41	44.6	133	US-09-252-991A-30896	Sequence 30896, A
23	41	44.6	177	US-09-252-991A-27363	Sequence 27363, A
24	41	44.6	204	US-09-252-991A-21317	Sequence 21317, A
25	41	44.6	226	US-09-252-991A-22052	Sequence 22052, A
26	41	44.6	255	US-09-252-991A-19598	Sequence 19598, A
27	41	44.6	404	US-09-489-039A-11606	Sequence 11606, A

28	41	44.6	478	4	US-09-252-991A-25411	Sequence 25411, A
29	41	44.6	482	4	US-09-252-991A-16654	Sequence 16654, A
30	41	44.6	686	4	US-09-252-991A-20509	Sequence 20509, A
31	40.5	44.0	96	4	US-09-252-991A-19216	Sequence 19216, A
32	40	43.5	134	4	US-09-252-991A-23430	Sequence 23430, A
33	40	43.5	210	4	US-09-134-001C-4065	Sequence 4065, Ap
34	40	43.5	213	4	US-09-252-991A-32093	Sequence 32093, A
35	40	43.5	237	4	US-09-252-991A-32304	Sequence 32304, A
36	40	43.5	256	4	US-09-252-991A-26244	Sequence 26244, A
37	40	43.5	421	4	US-09-252-991A-32336	Sequence 32336, A
38	40	43.5	442	4	US-09-252-991A-23554	Sequence 23554, A
39	40	43.5	646	4	US-09-328-352-6017	Sequence 6017, Ap
40	40	43.5	722	4	US-09-984-890-4	Sequence 4, Appl
41	40	43.5	724	4	US-09-984-890-2	Sequence 2, Appl
42	40	43.5	745	4	US-09-523-849-36	Sequence 36, Appl
43	40	43.5	877	4	US-09-252-991A-25547	Sequence 25547, A
44	40	43.5	908	1	US-08-356-354-6	Sequence 6, Appl
45	40	43.5	908	2	US-08-778-656-6	Sequence 6, Appl

ALIGNMENTS

```

RESULT 1
US-09-252-991A-22901
Sequence 22901, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22901
LENGTH: 629
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22901

Query Match      53.3%; Score 49; DB 4; Length 629;
Best Local Similarity 52.9%; Pred. No. 5.6;
Matches 9; Mismatches 4; Indels 0; Gaps 0;

QY      3 QESTGRSGRSGRSGS 19
DB      272 RQSHGRAGRAGRHGS 288

RESULT 2
US-09-252-991A-19857
Sequence 19857, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19857
LENGTH: 406
TYPE: PRT

```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19857

Query Match 51.1%; Score 47; DB 4; Length 406;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSXGRSGSGS 19
|:|||||:
DB 127 HGO\$NRGDRAGQGRHGA 144

RESULT 3
US-09-252-991A-21702
Sequence 21702, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21702
LENGTH: 395
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21702

Query Match 48.9%; Score 45; DB 4; Length 395;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTRGRSXGRSGRG 18
|||:|:
DB 251 SHOG\$BGRGDHSGRPG 268

RESULT 4
US-09-252-991A-18605
Sequence 18605, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18605
LENGTH: 778
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18605

Query Match 48.9%; Score 45; DB 4; Length 778;
Best Local Similarity 58.8%; Pred. No. 30;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 OESTRGRSXGRSGSGS 19
|:|||||:
DB 11 QTATRG\$EGR\$RSGN 27

RESULT 5
US-09-252-991A-20245
Sequence 20245, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20245
LENGTH: 369
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20245

Query Match 47.8%; Score 44; DB 4; Length 369;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSXGRSGRG 18
|:|||||:
DB 221 HRRTRG\$AAGQGRHG 237

RESULT 6
US-09-252-991A-33040
Sequence 33040, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 33040
LENGTH: 374
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33040

Query Match 47.8%; Score 44; DB 4; Length 374;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 ESTRGRSXGRSGR 16
|:|||||:
DB 258 ETLRGRTAGRGR 270

RESULT 7
US-09-252-991A-19015
Sequence 19015, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-19015

```

```

Query Match          47.8%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      2 HQESTGRSGRSG 18
Db      94 HAEQDPGAGTGRAGRAG 110

```

```

RESULT 8
US-09-252-991A-25967
; Sequence 25967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25967
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25967

```

```

Query Match          47.8%; Score 44; DB 4; Length 518;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      8 GRSGRSGRSG 18
Db      415 GRSGRSGRSG 425

```

```

RESULT 9
US-09-297-981-22
; Sequence 22, Application US/09297981
; Patent No. 6362007
; GENERAL INFORMATION:
; APPLICANT: MEHENS, LYDIE
; APPLICANT: LUHMANN, REINHARD GEORGE
; APPLICANT: UNION, ANN
; APPLICANT: RAYNACKERS, JOSEPH
; TITLE OF INVENTION: METHYLATED, SMD HOMOLOGOUS PEPTIDES, REACTIVE WITH THE
; TITLE OF INVENTION: ANTIBODIES FROM SERA OF LIVING BEINGS AFFECTED WITH
; TITLE OF INVENTION: SYSTEMIC LUPUS ERYTHEMATOSUS
; FILE REFERENCE: INNS011--
; CURRENT APPLICATION NUMBER: US/09/297,981
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-297-981-22

```

```

Query Match          46.7%; Score 43; DB 4; Length 26;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      7 RGRSGRSGRSG 18
Db      7 RGRSGRSGRSG 18

```

```

RESULT 10
US-09-297-981-20
; Sequence 20, Application US/09297981
; Patent No. 6362007
; GENERAL INFORMATION:
; APPLICANT: MEHENS, LYDIE
; APPLICANT: LUHMANN, REINHARD GEORGE
; APPLICANT: UNION, ANN
; APPLICANT: RAYNACKERS, JOSEPH
; TITLE OF INVENTION: METHYLATED, SMD HOMOLOGOUS PEPTIDES, REACTIVE WITH THE
; TITLE OF INVENTION: ANTIBODIES FROM SERA OF LIVING BEINGS AFFECTED WITH
; FILE REFERENCE: INNS011--
; CURRENT APPLICATION NUMBER: US/09/297,981
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-297-981-20

```

```

Query Match          46.7%; Score 43; DB 4; Length 46;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      7 RGRSGRSGRSG 18
Db      27 RGRSGRSGRSG 38

```

```

RESULT 11
US-09-252-991A-27107
; Sequence 27107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27107
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27107

```

```

Query Match          46.7%; Score 43; DB 4; Length 159;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY	1	SHQESTRGRSXGRSGRSGS	19
Dp	5	SASSSTRGRSAWRSTASSS	23

```

RESULT 12
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513.057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PR1
; ORGANISM: Cardamine oligosperma
US-09-513-057C-13

```

Query Match	48.7%;	Score 43;	DB 4;	Length 653;
Best Local Similarity	52.6%;	Pred. No. 51;		
Matches 10;	Conservative	1;	Mismatches 8;	Indels 0;
				Gaps 0;

```
QY      1 SHQESTRGRSXGRSGRSGS 19
          | | | | |
          : | | | |
Db      573 SRQVSTASSASGREGISGS 591
```

```

RESULT 13
US-09-252-991A-29339
: Sequence 29339, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 29339
: LENGTH: 85
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29339

```

Query Match	45.7%	Score 42;	DB 4;	Length 85;
Best Local Similarity	80.0%	Pred. No. 9;		
Best Match	8;	Conservative	1;	Indels 0; Gaps 0;

QY	8	GRSXGRSGRS	17
		:	
Db	73	GRAAGRSGRS	82

RESULT 14
US-09-252-991A-32394
; Sequence 32394, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

```

? CURRENT FILING DATE: 1999-02-18
? PRIOS APPLICATION NUMBER: US 60/074,788
? PRIOS FILING DATE: 1998-02-18
? PRIOS APPLICATION NUMBER: US 60/094,190
? PRIOS FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 33394
? LENGTH: 105
? TYPE: PRF
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33394

```

Query Match	45.7%;	Score 42;	DB 4;	Length 105;
Best Local Similarity	61.5%;	Pred. No. 11;		
Matches	8;	Conservative	1;	Mismatches 4;
				Indels 0;
				Gaps 0;

QY 6 TRGRSXGRSGRG 18
||| | :
Db 65 TRGRCAGRCGATG 77

```

RESULT 15
US-09-252-991A-21788
: Sequence 21788, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 21788
: LENGTH: 175
: TYPE: PR1
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21788

```

Query Match	45.7%;	Score 42;	DB 4;	Length 175;
Best Local Similarity	50.0%;	Pred. No. 19;		
Matches	8;	Conservative	3;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      1 SHQESTGRGSXGRGR 16
        | : ||| : |||
Db      62 SPRPARGRAAGRTGR 77
```

Search completed: September 28, 2004, 06:26:42
Job time : 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-4
Perfect score: 92
Sequence: 1 SHOESTRGRSGSGSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	2248	2 A35938	profilaggrin - hum
2	84	91.3	416	2 A32947	profilaggrin precursor
3	75	81.5	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2M47	E2 protein - human
5	50	54.3	800	2 T02852	probable membrane
6	48	52.2	772	2 T27512	hypothetical prote
7	48	52.2	822	2 T51049	related to nucleol
8	47	51.1	471	2 T33997	hypothetical prote
9	46	50.0	373	2 T02976	probable DNA bindi
10	46	50.0	836	2 G84727	probable DNA topoi
11	45	50.0	1232	2 S40766	hypothetical prote
12	46	50.0	1829	2 T35681	probable sensory h
13	45	49.5	135	2 A46338	RNA-binding protei
14	44.5	48.4	553	2 T27245	hypothetical prote
15	44.5	48.4	568	2 H88904	protein Y57G11C.9
16	44.5	48.4	659	2 T27246	hypothetical prote
17	44	47.8	229	2 T07219	nuclear protein SR
18	44	47.8	776	2 A87330	sensory box histid
19	44	47.8	790	2 T05576	hypothetical prote
20	44	47.8	867	2 T27135	hypothetical prote
21	44	47.8	871	2 T27135	hypothetical prote
22	44	47.8	2420	2 A84652	hypothetical prote
23	43.5	47.3	463	2 T51194	hypothetical prote
24	43.5	47.3	2561	2 T24864	hypothetical prote
25	43	46.7	123	2 T16234	hypothetical prote
26	43	46.7	150	2 B71185	probable ribosoma
27	43	46.7	151	2 F75145	lsu ribosomal prot
28	43	46.7	210	2 C84404	hypothetical prote
29	43	46.7	521	2 B86332	hypothetical prote

30	43	46.7	638	2 T53169	cytochrome 2 - hu
31	43	46.7	1507	2 B47328	natural killer cel
32	42	45.7	231	2 U70463	nuclear factor I -
33	42	45.7	306	2 T21220	hypothetical prote
34	42	45.7	350	2 A40459	nuclear phosphopro
35	42	45.7	369	2 F96788	protein T4012.22 [
36	42	45.7	374	2 A37282	52K active chroma
37	42	45.7	538	2 H86335	T20H2.2 protein -
38	42	45.7	694	2 S71786	wingless receptor
39	42	45.7	849	2 A96592	hypothetical prote
40	42	45.7	1015	2 T42013	frequency clock pr
41	42	45.7	1337	2 T30291	dextranase - strep
42	41.5	45.1	483	2 S36470	E2 protein - human
43	41	44.6	136	2 T35632	probable transposa
44	41	44.6	174	2 A27293	probable Y chromos
45	41	44.6	176	2 F71370	probable single-st

ALIGNMENTS

RESULT 1

A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gam: S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; PMID:91064347; PMID:2246957
A/Accession: A35938
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-References: GB:J02929
C/Genetics:
A/Gene: GDB:FLG
A/Cross-References: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
F/46-569/Region: profilaggrin repeat
F/570-893/Region: profilaggrin repeat
F/1074-1397/Region: profilaggrin repeat
F/1573-1896/Region: profilaggrin repeat

Query Match 97.8%; Score 90; DB 2; Length 2248;
Best Local Similarity 94.7%; Pred. No. 8.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGSGSGS 19
DB 227 SHOESTRGRSGSGSGS 245

RESULT 2

A32947
profilaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinley-Grant, D.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human profilaggrin and localization of th
A/Reference number: A32947; PMID:89269901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-References: GB:M24355; NID:g182604; PID:AAA52454.1; PID:g182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-q21
A:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 91.3%; Score 84; DB 2; Length 416;
Best Local Similarity 89.5%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGS 19
DB 7 SHOESTRGRSGRSGS 25

RESULT 3

A:Accession: A45135
A:Molecule type: DNA
A:Protein: profilaggrin - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A45135
R:Freeland, R.B.; Haydock, P.V.; Fleckman, P.; Nirmalskirti, W.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992
A>Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A:Reference number: A45135; MUID:93054736; PMID:1429717
A:Accession: A45135
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-591 <PRE>
A:Cross-references: GB:L01089; GB:M90567; NID:g190408; PIDN:AAA60177.1; PID:9553621
A:Note: sequence extracted from NCBI backbone (NCBI:P:118773)
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 81.5%; Score 75; DB 2; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.00051;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGS 19
DB 449 SHOESTRGRSGRSGS 467

RESULT 4

A:Accession: M2ML47
A:Molecule type: human papillomavirus type 47
C:Species: human papillomavirus type 47
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: D35324
R:Kiyo, T.; Adachi, A.; Ishibashi, M.
Virology 177, 401-405, 1990
A>Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
A:Reference number: A45324; MUID:90281611; PMID:2162112
A:Accession: D35324
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-506 <KIT>
A:Cross-references: GB:M2305; NID:g333062; PIDN:AAA4679.1; PID:g333067
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.19;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGS 19
DB 1 SHOESTRGRSGRSGS 19

DB 342 SRENTGRGRGRGRGAGS 360

RESULT 5
T02852
probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)

C:Species: Leishmania major
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: T02852; H81462
R:Wyle, P.J.
submitted to the EMBL Data Library, May 1998
A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
A:Reference number: Z14740
A:Accession: T02852
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-800 <MYL>
A:Cross-references: EMBL:AE001274; NID:g3264850; PID:g2266920
R:Wyle, P.J.; Audleman, L.; Devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A>Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: H81462
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-800 <PYL>
A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AC24675.1; PID:g2266920; GSPDB:GN001
C:Experimental source: strain MOW/IL/81/Friedlin
C:Genetics:
A:Gene: L1439.4
A:Map position: 1

Query Match 54.3%; Score 50; DB 2; Length 800;
Best Local Similarity 55.6%; Pred. No. 5.2;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTRGRSGRSGS 19
DB 429 HRDGVGRGLSTASGRSGS 446

RESULT 6

A:Accession: T27512
A:Molecule type: hypothetical protein ZC302.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2003
C:Accession: T27512
R:Kelly, P.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z20380
A:Accession: T27512
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-772 <WIL>
A:Cross-references: EMBL:Z73978; PIDN:CA98292.1; GSPDB:GN00023; CESP:ZC302.1
C:Experimental source: clone ZC302
C:Genetics:
A:Gene: CESP:ZC302.1
A:Map position: 5
A:Intons: 43/2; 70/3; 94/2; 129/2; 393/3; 548/3; 558/3; 642/3
C:Superfamily: double-strand break repair protein MRE11; phosphatase core homology

Query Match 52.2%; Score 48; DB 2; Length 772;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGS 19
DB 697 SKQPTTRGGRGARGAGAS 715

RESULT 7

T51049

related to nucleolar phosphoprotein [imported] - Neurospora crassa
N:Alternate names: protein B12F1.10
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51049
R:Schulte, U.; Allyn, V.; Hohenfeld, U.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z5286
A:Accession: T51049
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-822 <SCH>
A:Cross-references: EMBL:AJ390091; GSPDB:GN00116; NCSP:B12F1.10
A:Experimental source: BAC clone B12F1; strain OR74A
C:Genetics:
A:Gene: NCSP:B12F1.10
A:Map position: 6
A:Introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
Best Local Similarity 52.8%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGSRGXSGRSGS 19
: ||| ||| ||| |||
Db 414 RESASGRTRGRGRGRT 430

RESULT 8

hypothetical protein W03G1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33997
R:Pauley, A.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: EMBL:AF125664; PIDN:AA14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
Best Local Similarity 52.6%; Pred. No. 9.2;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGS 19
: ||| ||| ||| |||
Db 167 SSRSPSRGRGRGGRSGS 185

RESULT 9

probable DNA binding protein PCF2 - rice
C:Species: Oryza sativa (rice)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T02976
R:Kosugi, S.; Ohashi, Y.
Plant Cell 9, 1607-1619, 1997
A:Title: PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell
A:Reference number: Z14803; MUID:91480096; PMID:9338963
A:Accession: T02976
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-373 <KOS>
A:Cross-references: EMBL:DB9261; NID:g2580439; PIDN:BAA23143.1; PID:g2580440
A:Experimental source: cultivar Nipponbare

Query Match 50.0%; Score 46; DB 2; Length 373;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGS 15
||| ||| ||| |||
Db 334 SHEQRRGRGRKEGNSG 348

RESULT 10

G84727
probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84727
R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
enes, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617127
A:Accession: G84727
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <STO>
A:Cross-references: GB:AE002093; NID:g4263718; PIDN:AA15404.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32000
A:Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSRGXSGRSGS 19
||| ||| ||| |||
Db 812 RGRGRGGRGGRSGS 824

RESULT 11

hypothetical protein ZK512.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Sep-1997
C:Accession: S40766
R:Hawkins, T.; Ainscough, R.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40759
A:Accession: S40766
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1232 <HAM>
A:Cross-references: EMBL:Z22177; NID:g297989; PID:g297994
C:Genetics:
A:Introns: 5/3; 19/1; 52/2; 108/2; 306/3; 344/2; 490/3; 542/3; 634/3; 703/2; 1060/3; 1110

Query Match 50.0%; Score 46; DB 2; Length 1232;
Best Local Similarity 64.3%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HOESTRGSRGXSG 15
||| ||| ||| |||
Db 110 HONSSRGSRGFRSG 123

RESULT 12

T35681
probable sensory histidine kinase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35681
R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Raftery, M.A.

submitted to the EMBL Data Library, July 1998
A:Reference number: Z21587

A:Accession: T35681

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1829 <BAR>

A:Cross-references: EMBL:AL031031; PIN:CAI19849.1; GSPDB:GN00070; SCOEDB:SC7C7.03

A:Experimental source: strain A3(2)

A:Genetics:

A:Gene: SCOEDB:SC7C7.03

Query Match 50.0%; Score 46; DB 2; Length 1829;
Best Local Similarity 60.0%; Pred. No. 49;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 STRGRSGXSGRSGS 19
| : ||| ||| |||
| : ||| ||| |||
DB 21 SNGSRSGRSGRSGS 35

RESULT 13

A46398

RNA-binding protein 1 - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 24-Sep-1999

C:Accession: A46398

R:Kim, Y.J.; Zuo, P.; Manley, J.L.; Baker, B.S.

Genes Dev. 6, 2569-2579, 1992

A:Title: The *Drosophila* RNA-binding protein RBP1 is localized to transcriptionally active

A:Reference number: A46398; MUID:94040720; PMID:1340470

A:Accession: A46398

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-135 <KIM>

A:Cross-references: GB:L04929; NID:G158223; PID:AAA28850.1; PID:G158224

A>Note: authors translated the codon GCT for residue 65 as Arg

C:Genetics: rbp1

A:Gene: FlyBase:Rbp1

A:Cross-references: FlyBase:FBgn0010252

A:Introns: 106/2

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein

F:12-74/Domain: ribonucleoprotein repeat homology <RKM>

Query Match 49.5%; Score 45.5; DB 2; Length 135;
Best Local Similarity 52.2%; Pred. No. 4.7;

Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 4 ESTGRSR-----GXSGRSGS 19
| : ||| ||| |||
| : ||| ||| |||
DB 80 EMSGSRDRRRGEGSGRSGS 102

RESULT 14

T27245

hypothetical protein Y57G11C.9a - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27245

R:McMurray, A.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27245

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-553 <ML>

A:Cross-references: EMBL:Z99281; PID:CA854457.1; GSPDB:GN00022; CESP:Y57G11C.9a

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.9a

A:Map position: 4

A:Introns: 67/2; 109/2; 382/1; 418/2; 447/1; 496/3

Query Match 48.4%; Score 44.5; DB 2; Length 553;

Best Local Similarity 64.7%; Pred. No. 26;
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 CESTGRSGXSGRSGS 19
| : ||| ||| |||
| : ||| ||| |||
DB 218 RSTGRSR-SGGRSGS 233

RESULT 15

H88904

protein Y57G11C.9 [imported] - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: H88904

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.eleg

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: H88904

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <STO>

A:Cross-references: GB:chr_IV; PID:CA816531.1; PID:G3881206; GSPDB:GN00022; CESP:Y57G11C

A:Gene: Y57G11C.9

A:Map position: 4

Query Match 48.4%; Score 44.5; DB 2; Length 568;
Best Local Similarity 64.7%; Pred. No. 27;

Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 CESTGRSGXSGRSGS 19
| : ||| ||| |||
| : ||| ||| |||
DB 233 RSTGRSR-SGGRSGS 248

Search completed: September 28, 2004, 06:14:59
Job time : 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)

136,952 Million cell updates/sec

Title: US-09-308-150-4

Sequence: 1 SHOESTRGSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	91.3	416	1	P20930 homo sapien
2	58	63.0	506	1	P22420 human papil
3	48	52.2	772	1	Q23255 caenorhabd
4	46	50.0	963	1	Q14157 homo sapien
5	46	50.0	1232	1	P34603 caenorhabd
6	45.5	49.5	135	1	Q02427 drosophila
7	44	47.8	124	1	Q9V002 drosophila
8	43	46.7	123	1	Q19952 caenorhabd
9	43	46.7	150	1	Q59437 pyrococcus
10	43	46.7	151	1	Q9V433 pyrococcus
11	43	46.7	210	1	Q9H117 mus musculu
12	43	46.7	638	1	Q01546 homo sapien
13	43	46.7	1453	1	P30445 mus musculu
14	42	45.7	231	1	P14057 sus scrofa
15	42	45.7	375	1	P26666 drosophila
16	42	45.7	694	1	Q9V433 drosophila
17	42	45.7	1015	1	Q9V433 drosophila
18	42	45.7	1180	1	Q9H117 mus musculu
19	42	45.7	1337	1	P39653 streptococ
20	41.5	45.1	388	1	P38159 mus musculu
21	41.5	45.1	391	1	P38159 mus musculu
22	41.5	45.1	483	1	P36783 human papil
23	41	44.6	174	1	P09102 homo sapien
24	41	44.6	176	1	P09102 homo sapien
25	41	44.6	205	1	P28007 saccharomy
26	41	44.6	414	1	P28007 saccharomy
27	41	44.6	454	1	P28007 saccharomy
28	41	44.6	502	1	P28007 saccharomy
29	41	44.6	503	1	P50809 human papil
30	41	44.6	509	1	P50809 human papil
31	41	44.6	713	1	P64060 rattus norv
32	41	44.6	825	1	P63003 rattus norv
33	41	44.6	2404	1	Q9G447 mus musculu

ALIGNMENTS

RESULT 1	ID	FILE HUMAN	STANDARD	PRT	416 AA
AC	P20930	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Flaggrin precursor (Fragment).				
GN	Flag.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89296901; PubMed=2740331;				
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,				
RA	Carnizzaro L., Croce C.W., Huebner K., Lessin S.R., Steinhert P.M.,				
RT	"Characterization of a cDNA clone encoding human flaggrin and				
RT	localization of the gene to chromosome region 1q21."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).				
RN	[2]				
RP	CITRULLINATION.				
RX	MEDLINE=96374388; PubMed=8780679;				
RA	Senhhu T., Kan S., Ogawa H., Manabe W., Asaga H.,				
RT	"Preferential deimination of keratin Ki and flaggrin during the				
RT	terminal differentiation of human epidermis."				
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).				
CC	-I- FUNCTION: Aggregates keratin intermediate filaments and promotes				
CC	disulfide-bond formation among the intermediate filaments during				
CC	terminal differentiation of mammalian epidermis.				
CC	-I- PFM: Flaggrin is initially synthesized as a large, insoluble,				
CC	highly phosphorylated precursor containing many tandem copies of				
CC	324 AA, which are not separated by "large linker". The precursor				
CC	is deposited as keratohyalin granules. During terminal				
CC	differentiation it is dephosphorylated and proteolytically				
CC	cleaved.				
CC	-I- PFM: Undergoes deimination of some arginine residues				
CC	(citrullination).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL: M24355; AA52454.1; -				
CC	PIR: A32947; A32947.				
CC	GeneW: HGNC:3748; FLG.				
CC	-----				
CC	DR GO:0005882; C:intermediate filament; NAS.				
CC	DR GO:0005198; F:structural molecule activity; NAS.				
CC	DR GO:0007275; P:development; NAS.				
CC	DR InterPro: IPR003303; Flaggrin.				

34	41	44.6	2426	1	SON_HUMAN	P18583 homo sapien
35	40.5	44.0	1509	1	MYSN_ACACA	P05659 acanthamoeb
36	40	43.5	150	1	YDPC_SCHPO	O14015 schizosacch
37	40	43.5	356	1	DPF3_MOUSE	P58269 mus musculu
38	40	43.5	526	1	2ABA_YEAST	O00362 saccharomy
39	40	43.5	546	1	GHT5_SCHPO	P78831 schizosacch
40	40	43.5	659	1	HMT1_DROME	P22807 drosophila
41	40	43.5	677	1	SG1_MOUSE	P16014 mus musculu
42	40	43.5	730	1	DPOL_HPBVA	P12933 hepatitis b
43	40	43.5	842	1	DPOL_HPBVA	P31870 hepatitis b
44	40	43.5	843	1	DPOL_HPBVA	P03157 hepatitis b
45	40	43.5	1053	1	SPS_SOLTU	Q43845 solanum tub

DR Pfam, PF03516; Flaggrin; 2.
 DR PRINTS; PR00487; FLAGGRIN.
 KW Phosphorylation; Cytullination; Developmental protein.
 FT NON TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
 Best Local Similarity 89.5%; Pred. No. 4.4e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGRSGS 19
 DB 7 SHOESTRGRSGXSGRSGS 25

RESULT 2
 ID VE2_HPVA7 STANDARD; PRT; 506 AA.
 AC P22420;

DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.

OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC Papillomavirus.

OX NCBI_TaxID=10594;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 type 47 inferred from its DNA sequence.";
 RU Virology 177:401-405 (1990).

CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION
 IT BINDS TO THE E2B RESPONSE ELEMENT (5'-ACNNNNNGT-3') PRESENT
 IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2B/S POSITION
 WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 REPLICATION.

CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; M32305; AAA46979.1; -.
 DR PIR; D35324; W2WT47.
 DR HSSP; P03122; 2BOP.
 DR InterPro; IPR000427; E2_C.
 DR InterPro; IPR001866; E2_N.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF00511; E2_C; 1.
 DR Pfam; PF00506; E2_N; 1.
 DR Pfam; PF000672; E2_C; 1.
 DR Pfam; PF000678; E2_N; 1.
 DR ProDom; PD000672; E2_C; 1.
 DR ProDom; PD000678; E2_N; 1.
 DR Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BFB755065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.07;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGRSGS 19

DB 342 SREGNTRGRGRGRGRS 360

RESULT 3
 ID MR11_CAEEL STANDARD; PRT; 772 AA.
 AC Q23255;

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Double-strand break repair protein mre-11.
 GN MRE-11 OR ZC302.1.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.; FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
 RP GLU-139.
 RX MEDLINE=21135651; PubMed=11238374;
 RA Chin G.M., Villeneuve A.M.;

RT "C. elegans mre-11 is required for meiotic recombination and DNA
 repair but is dispensable for the meiotic G(2) DNA damage
 checkpoint.";
 RT Genes Dev. 15:522-534 (2001).

RL Genes Dev. 15:522-534 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;

RA Kelly P.F.;
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
 CC processing.

CC -1- FUNCTION: Required for meiotic crossing over and chiasma
 CC formation. Pachytene morphology and homolog pairing are normal.
 CC Vital in long term for maintenance of reproductive capacity of

CC subsequent generations.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: Belongs to the MR11/RAD32 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; Z73978; CA98292.1; -.
 DR PIR; T27512; T27512.
 DR GenomOnline; 208612; -.
 DR WormPep; ZC302.1; CE06573.
 DR InterPro; IPR003701; DNA_repair.
 DR InterPro; IPR004643; M-Pdestrase.
 DR InterPro; IPR007281; Mre11_DNA_bind.
 DR Pfam; PF00149; Metad10pos; 1.
 DR Pfam; PF04152; Mre11_DNA_bind; 1.
 DR TIGRFams; TIGR00583; mre11; 1.
 KW DNA repair; Hydrolyase; Nuclease; Endonuclease; Exonuclease;
 KW Nuclear protein; Manganese; Meiosis.
 FT MUTAGEN 139 139
 FT E->K: IN MRE-11-ME41; DEFECTIVE IN
 MEIOTIC CHROMOSOME DEGRADATION.
 SQ SEQUENCE 772 AA; 86913 MW; 744A0754C2604C4B CRC64;

Query Match 52.2%; Score 48; DB 1; Length 772;
 Best Local Similarity 52.6%; Pred. No. 4.2;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGRSGS 19

```

Db      697 SKOPTTRGRGRGARGAGAS 715

RESULT 4
Y144_HUMAN
ID      Y144_HUMAN      STANDARD;          PRT;          983 AA.
AC      014157;
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Hypothetical protein KIAA0144.
GN      KIAA0144.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Bone marrow;
RX      MEDLINE=96127530; PubMed=8590280;
RA      Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT      "Prediction of the coding sequences of unidentified human genes. IV.
RT      The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT      analysis of cDNA clones from human cell line KG-1."
RL      DNA Res. 2:167-174(1995).
CC      -1- SIMILARITY: Contains 1 UBA domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D63478; BAA09765.1; -.
DR      InterPro; IPR000449; UBA_domain.
DR      Pfam; PF00627; UBA; 1.
DR      SMART; SM00165; UBA; 1.
DR      PROSITE; PSS0030; UBA; 1.
KM      Hypothetical protein.
FT      DOMAIN 49
FT      UBA.
SQ      SEQUENCE 983 AA; 103930 MW; 50579CFA293EEA68 CRC64;

Query Match      50.0%; Score 46; DB 1; Length 983;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy      7 RGRSGRGSGSG 18
Db      177 RGRGRGSGSGRRG 188

RESULT 5
Y005_CAEEL
ID      Y005_CAEEL      STANDARD;          PRT;          1232 AA.
AC      P34643;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical protein ZK512.5 in chromosome III.
GN      ZK512.5.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxId=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RX      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Betks M.,
RA      Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

```

```

RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA      Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA      Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA      Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callahan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders B., Showkeen R.,
RA      Sims M., Smaildon N., Smith A., Smith M., Sonnenauer B., Staden R.,
RA      Sulten J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA      Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA      Woldman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans."
RL      Nature 368:32-38(1994).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; 222177; CAA80146.1; -.
DR      PIR; S40766; S40766.
DR      MornPeP; ZK512.5; CE00411.
KM      Hypothetical protein.
SQ      SEQUENCE 1232 AA; 134923 MW; 6DFC35D664A8D6A CRC64;

Query Match      50.0%; Score 46; DB 1; Length 1232;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      2 HOESTRGSRGSG 15
Db      110 HONSSRGPSGSPSG 123

RESULT 6
RBP1_DROME
ID      RBP1_DROME      STANDARD;          PRT;          135 AA.
AC      Q02427; Q26271; Q9VGM8;
DT      15-DEC-1998 (Rel. 37, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      RNA-binding protein 1.
GN      RBP1 OR CG37136.
OS      Drosophila melanogaster (Fruit Fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxId=7227;
RN      [1]
RP      SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX      MEDLINE=94040720; PubMed=1340470;
RA      Kim Y.-J., Zuo P., Wanley J.L., Baker B.S.;
RT      "The Drosophila RNA-binding protein RBP1 is localized to
RT      transcriptionally active sites of chromosomes and shows a functional
RT      similarity to human splicing factor ASF/SF2."
RL      Genes Dev. 6:2563-2579(1992).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RC      STRAIN=Berkley;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blaise R.G., Change M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKis G.L.G.,
RA      Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA      Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

```

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 14-52 FROM N.A.
 RX MEDLINE=93109300; PubMed=8417324;
 RA Kim Y.J., Baker B.S.;
 RT "Isolation of Ruv-type RNA-binding protein genes and the analysis of
 RL their relatedness by using a numerical approach,"
 RN Mol. Cell. Biol. 13:174-183(1993).
 RN [4]
 RP FUNCTION:
 RX MEDLINE=95393975; PubMed=7664738;
 RA Heinrichs V., Baker B.S.;
 RT "The *Drosophila* SR protein RBP1 contributes to the regulation of
 RT doublesex alternative splicing by recognizing RBP1 RNA target
 RT sequences,"
 RL EMO U. 14:3987-4000(1995).
 CC -1- FUNCTION: Contributes to the activation of female-specific DSX
 CC splicing in vivo by recognizing the RBP1 target sequences within
 CC the purine-rich polypyrimidine tract of the female-specific 3'
 CC splice site.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RBP1-A;
 CC IsoId=Q02427-1; Sequence=Displayed;
 CC Name=2; Synonyms=RBP1-B;
 CC IsoId=Q02427-2; Sequence=VSP_005817;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
 CC -1- PTM: Extensively phosphorylated on serine residues in the RS
 CC domain (Probable).
 CC -1- SIMILARITY: Belongs to the splicing factor SR family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.jsb-sib.ch/announce/>
 CC or send an email to license@jsb-sib.ch).
 CC -----
 DR EMBL, L04829, AAA28850.1, -;
 DR EMBL, AE003688, AAF54555.1, -;
 DR EMBL, S51691, AAB24622.1, -;
 DR PIR, A46398, A46398.
 DR PIR, A48110, A48110.

DR flyBase; FBgn0010252; Rbp1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00706; rrm; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PSS0030; RRM_RNP_1; FALSE_NEG.
 DR Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
 KW Phosphorylation.
 FT DOMAIN 11 84 RNA-BINDING (RRM).
 FT DOMAIN 82 135 ARG/SER-RICH (RS DOMAIN).
 FT VARSPIC 107 135 Missing (in isoform 2).
 FT VARSPIC 107 135 /FTid=VSP_005817.
 FT CONFLICT 14 14 Y -> F (IN REF. 3).
 FT CONFLICT 65 65 R -> A (IN REF. 1).
 SQ SEQUENCE 135 AA; 1546 MW; 895DCE902518D991 CRC64;
 Query Match 49.5%; Score 45.5; DE 1; Length 135;
 Best local Similarity 52.2%; Pred. No. 1.6;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
 Db 80 EMSGRSRRRRRGEGSGSGSGS 102
 QY 4 ESTGRSR-----GXSGRSGS 19
 ID SMD1_DROME STANDARD; PRT; 124 AA.
 AC Q9V002; Q8SYR6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable small nuclear ribonucleoprotein Sm D1 (snRNP core protein D1)
 DE (Sm-D1)
 OS SNRNP69D OR ECEDNA:RE39488 OR C010753.
 GN *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blake J.R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

```

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley S.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=2242606; Pubmed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Garth H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- FUNCTION: Essential for pre-mRNA splicing. Implicated in the
CC formation of stable, biologically active snRNP structures (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the snRNP core protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003540; AAF49893.1; -.
DR EMBL: AY071359; AAU48981.1; -.
DR FLYbase; FBgn0016940; snRNP69D.
DR GO; GO:0030532; C:small nuclear ribonucleoprotein complex; ISS.
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; ISS.
DR GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; ISS.
DR InterPro: IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR Nucleic Acid; Nucleic Acid; Ribonucleoprotein; mRNA splicing; mRNA processing.
FT DOMAIN 86 124 ARG/LYS-RICH (BASIC).
FT DOMAIN 100 118 9 X 2 AA TANDDEM REPEATS OF R-G.
FT CONFLICT 52 52 P -> H (IN REF. 2).
SQ SEQUENCE 124 AA; 13796 MW; 36FA51D3FB68E5 CRC64;
Query Match 47.8%; Score 44; DB 1; Length 124;
Best Local Similarity 50.0%; Pred. No. 2.6;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
Oy 1 SHOESTRGRSGRGSG 18
Db 94 SGRVGRGRGRGRGRGG 111

```

```

CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20864; AAC46661.1; -.
DR PIR; T16234; T16234.
DR Wormpep; F32A5.7; CE01277.
DR InterPro; IPR006549; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
DR Nucleic Acid; Nucleic Acid; Ribonucleoprotein; mRNA splicing; mRNA processing;
KW RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;
Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Oy 3 QESTRGRSGRGSG 18
Db 93 RQSRGRSGRGSG 108

```

```

RESULT 9
RL19 PYRHO STANDARD; PRT; 150 AA.
AC 059437;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L19B.
GN RPL19B OR PH1759.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OC NCB1_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; Pubmed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- SIMILARITY: Belongs to the L19B family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000007; BAA30873.1; -.

```

DR PIR; B71185; Ribosomal_L19e.
 DR InterPro: IPR00196; Ribosomal_L19e.
 DR Pfam; PF01280; Ribosomal_L19e; 1.
 DR ProDom; PD004823; Ribosomal_L19e; 1.
 DR PROSITE; PS00526; RIBOSOMAL_L19E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 150 AA; 17997 MW; 412223A854825B69 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 150;
 Best Local Similarity 41.2%; Pred. No. 4.5;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSRGXSGRSG 18
 DB 67 HEOKKGRHRGPGSRKG 83

RESULT 10

RL19_PYPAB STANDARD; PRT; 151 AA.
 AC Q9V1V3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 50S ribosomal protein L19S.
 GN RPL19E OR PYPAB03230 OR PAB2134.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CC NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RX MEDLINE=22511545; PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
 RA Poch O., Priet D., Querellou J., Ripp R., Thierry J.-C.,
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P;
 RT "An integrated analysis of the genome of the hyperthermophilic
 archaeon Pyrococcus abyssi."
 RL Mol. Microbiol. 47:1493-1512(2003).
 CC -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ248284; CAB49245.1; -
 DR PIR; F75145; F75145.
 DR InterPro; IPR00196; Ribosomal_L19e.
 DR Pfam; PF01280; Ribosomal_L19e; 1.
 DR ProDom; PD004823; Ribosomal_L19e; 1.
 DR PROSITE; PS00526; RIBOSOMAL_L19E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 151 AA; 18187 MW; 85FC056443437A0 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 151;
 Best Local Similarity 41.2%; Pred. No. 4.5;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSRGXSGRSG 18
 DB 67 HEOKKGRHRGPGSRKG 83

RESULT 11

YO43_HALN1 STANDARD; PRT; 210 AA.
 AC Q9HMH2;
 DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein Vng2543c.
 GN VNG2543C.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 CC Halobacteriaceae; Halobacterium.
 CC NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahatras G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angwine C.M., Dale H.,
 RA Tsenbarger T.A., Beck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RA "Genome sequence of Halobacterium species NRC-1."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- SIMILARITY: Contains 1 AMMECR1 domain.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE005130; AAG20599.1; -
 DR PIR; C84404; C84404.
 DR HAMAP; MF_00645; atypical; 1.
 DR InterPro; IPR002733; DUF51.
 DR Pfam; PF01871; AMMECR1; 1.
 DR ProDom; PD009671; DUF51; 1.
 DR TIGRfam; TIGR00296; TIGR00296; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 210 AA; 22964 MW; D4CB0CBE8307EFF CRC64;

Query Match 46.7%; Score 43; DB 1; Length 210;
 Best Local Similarity 61.1%; Pred. No. 6.5;
 Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 4 ESTRGRS--RGXSGRSGS 19
 DB 55 ESTRGRGRURCGAGAGS 72

RESULT 12

K220_HUMAN STANDARD; PRT; 638 AA.
 ID K220_HUMAN
 AC Q01546;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2P) (K2P)
 DE (CK 2P).
 GN KRT2P.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93114504; PubMed=1282112;
 RA Collin C., Onayoun J.P., Grund C., Franke W.W.;
 RA "Suprabasal marker proteins distinguishing keratinizing squamous
 RT epithelia: cyokeratin 2 polypeptides of oral masticatory epithelium
 RT and epidermis are different."
 RL Differentiation 51:137-148(1992).
 CC -1- FUNCTION: Probably contributes to terminal cornification.

```

CC CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC CC keratinocytes.
CC CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC CC microfilillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC CC (neutral to basic; 56-70 kDa) [K1 to K8].
CC CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M99063; AAA35746.1; -.
CC CC PIR; I53169; I53169.
CC CC GO; GO:0005882; C:intermediate filament; NAS.
CC CC GO; GO:0005198; C:structural molecule activity; NAS.
CC CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
CC CC InterPro; IPR001684; IF.
CC CC InterPro; IPR002957; Keratin_I.
CC CC InterPro; IPR003054; Keratin_II.
CC CC Pfam; PF00038; filament; 1.
CC CC PRINTS; PR01248; TYPBKERATIN.
CC CC PRINTS; PR01276; TYPBKERATIN.
CC CC PROSITE; PS00226; IF; 1.
CC CC Intermediate filament; Coiled coil; Keratin; Phosphorylation.
CC CC DOMAIN 1 182 HEAD.
CC CC FT DOMAIN 183 492 ROD.
CC CC FT DOMAIN 493 638 TAIL.
CC CC FT DOMAIN 183 218 COIL_1A.
CC CC FT DOMAIN 219 237 LINKER_1.
CC CC FT DOMAIN 238 329 COIL_1B.
CC CC FT DOMAIN 330 353 LINKER_12.
CC CC FT DOMAIN 354 492 COIL_2.
CC CC FT MOD_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
CC CC SEQUENCE 638 AA; 65871 MW; 9B743AB8B72076AF CRC64;
CC CC -----
Query Match 46.7%; Score 43; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC CC -----
CC CC 3 QESTRGSRGSGSGS 17
CC CC ID NKRKR MOUSE STANDARD; PRT; 1453 AA.
CC CC AC P30415;
CC CC DT 01-APR-1993 (Rel. 25, Created)
CC CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC CC DE NK-tumor recognition protein (Natural-Killer cells cyclophilin-
CC CC DE related protein) (NK-TR protein).
CC CC GN NKTR.
CC CC OS Mus musculus (Mouse).
CC CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC CC NCBI_TaxID=10090;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RX MEDLINE=93133824; PubMed=8421688;
CC CC RA Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
CC CC RA Oretallo J.R.;
CC CC RT "A cyclophilin-related protein involved in the function of natural
CC CC RT killer cells.";
CC CC RL Proc. Natl. Acad. Sci. U.S.A. 90:542-546 (1993).
CC CC RN [2]
CC CC RP REVISIONS TO C-TERMINUS.

```

```

CC CC STRAIN=BA1B/c; TISSUE=Blood;
CC CC RA Anderson S.K.;
CC CC RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: Component of a putative tumor-recognition complex.
CC CC Involved in the function of NK cells.
CC CC -1- SIMILARITY: Contains 1 cyclophilin-like p1ase domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U04289; AAA37500.2; ALT_INIT.
CC CC DR HSSP; Q27450; IA33.
CC CC DR MGD; MGI:97346; Nktr.
CC CC DR InterPro; IPR002130; CSA_P1ase.
CC CC DR Pfam; PF00160; pro.isomerase; 1.
CC CC DR PRINTS; PR00153; CSAPRIMASE.
CC CC DR PROSITE; PR00170; CSA_P1ASE_1; 1.
CC CC DR PROSITE; PS50072; CSA_P1ASE_2; 1.
CC CC KW Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
CC CC FT DOMAIN 1 176 P1ASE, CYCLOPHILIN-TYPE.
CC CC FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 198 273 ARG/SER-RICH.
CC CC FT DOMAIN 468 565 ARG/SER-RICH.
CC CC FT DOMAIN 658 812 ARG/SER-RICH.
CC CC FT DOMAIN 1303 1453 ARG-SER TANDDEM REPEAT-RICH.
CC CC SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;
CC CC -----
Query Match 46.7%; Score 43; DB 1; Length 1453;
Best Local Similarity 81.8%; Pred. No. 51;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC CC -----
CC CC 9 RSRGSGSGSGS 19
CC CC ID NFIL_PIG STANDARD; PRT; 231 AA.
CC CC AC P14057;
CC CC DT 01-JAN-1990 (Rel. 13, Created)
CC CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC CC DE Nuclear factor 1 (NF-1) (CCAAT-box binding transcription factor) (CTF)
CC CC DE (TGGCA-binding protein) (Fragment).
CC CC GN NF1.
CC CC OS Sus scrofa (Pig).
CC CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
CC CC NCBI_TaxID=9823;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RX MEDLINE=88296883; PubMed=2841167;
CC CC RA Meisternernst M., Rogge L., Donath C., Gander I., Lottspeich F.,
CC CC RA Meitz R., Dobner T., Roessler R., Stelzer G., Winnacker E.L.;
CC CC RT "Isolation and characterization of the porcine nuclear factor I (NF1)
CC CC RT gene.";
CC CC RL FEBS Lett. 236:27-32 (1998).
CC CC RN [2]
CC CC RP PARTIAL SEQUENCE.
CC CC RC TISSUE=Liver;
CC CC RX MEDLINE=89088257; PubMed=3207762;
CC CC RA Gander I., Roessler R., Rogge L., Meisternernst M., Schneider R.,
CC CC RA Meitz R., Lottspeich F., Winnacker E.L.;
CC CC RT "Purification methods for the sequence-specific DNA-binding protein
CC CC RT nuclear factor I (NF1) -- generation of protein sequence

```

```

RT Information.
RL Biochim. Biophys. Acta 951:411-418(1988).
CC -1- FUNCTION: Recognizes and binds the palindromic sequence 5'-
CC TTGGCANNNGCCAA-3' present in viral and cellular promoters and in
CC the origin of replication of adenovirus type 2. These proteins are
CC individually capable of activating transcription and replication.
CC -1- SUBUNIT: Binds DNA as a homodimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=P14057-1; Sequence=Displayed;
CC -1- SIMILARITY: Belongs to the CTF/NF-I family.
CC -1- SIMILARITY: Contains 1 DNA/MH1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X12764; CAA31254.1; -.
CC DR PIR, J70463; J70463.
CC DR TRASNFPAC, T00173; -.
CC DR TRASNFPAC, T00538; -.
CC DR InterPro, IPR000647; CTF_NF1.
CC DR InterPro, IPR003619; DwaRtn_A.
CC DR Pfam, PF03165; MH1; 1.
CC DR SMART, SMO0523; DWA; 1.
CC DR PROSITE, PS00349; CTF_NF1; 1.
CC KW Transcription regulation; DNA replication; DNA-binding; Activator;
CC Nuclear protein; Multigene family; Alternative splicing.
CC FT DOMAIN 65 173 DWA.
CC FT NON_TER 231 231
CC SQ SEQUENCE 231 AA; 26162 MW; A463GBBE3DDE743 CRC64;
CC Query Match 45.7%; Score 42; DB 1; Length 231;
CC Best Local Similarity 53.3%; Pred. No. 10;
CC Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 4 ESTGRSRGXSGSG 18
Db 186 EAGRARGSGDREG 200

```

```

RC TISSUE=Embryo;
RX MEDLINE=91357476; PubMed=1885003;
RA Champlin D.T., Frasch M., Saunweber H., Lis J.T.;
RT "Characterization of a Drosophila protein associated with boundaries
RL of transcriptionally active chromatin."
RL Gene Dev. 5:1611-1621(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=92159037; PubMed=1741384;
RA Mayeda A., Zahler A.M., Krainer A.R., Roth M.B.;
RT "Two members of a conserved family of nuclear phosphoproteins are
RL involved in pre-mRNA splicing."
RL Proc. Natl. Acad. Sci. U.S.A. 89:1301-1304(1992).
RN [4]
RP FUNCTION.
RX MEDLINE=95021280; PubMed=7935465;
RA Ring H.Z., Lis J.T.;
RT "The SR protein B52/SRP5 is essential for Drosophila development."
RL Mol. Cell. Biol. 14:7499-7506(1994).
CC -1- FUNCTION: Essential for development. May have a critical role in
CC splicing or in controlling alternative splice site use of at least
CC some pre-mRNA in vivo. Not required for all splicing. May play a
CC general role in the condensation or decondensation of chromatin.
CC -1- SUBCELLULAR LOCATION: Nuclear; associated with boundaries of
CC transcriptionally active chromatin.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P26686-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P26686-2; Sequence=VSP_005878;
CC -1- DEVELOPMENTAL STAGE: Expressed throughout development.
CC -1- PTM: Extensively phosphorylated on serine residues in the RS
CC domain.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -1- SIMILARITY: Belongs to the SR family of splicing factors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X58720; CAA11556.1; -.
CC DR EMBL, X62599; CAA44483.1; -.
CC DR PIR, A40459; A40459.
CC DR FlyBase, FBgn0004587; B52.
CC DR GO, GO:001607; C:nuclear speck; IDA.
CC DR GO, GO:0005634; C:nucleus; IDA.
CC DR InterPro, IPR000504; RNA_rec_mot.
CC DR Pfam, PF00076; rrm; 2.
CC DR SMART, SMO0360; RRM; 2.
CC DR PROSITE, PS50102; RRM; 2.
CC DR PROSITE, PS00030; RRM_RNP_1; FALSE NEG.
CC KW Nuclear protein; Phosphorylation; mRNA splicing; RNA-binding;
CC Repeat; Alternative splicing.
CC FT INIT_MET 0 0
CC FT DOMAIN 3 73 RNA-BINDING (RRM) 1.
CC FT DOMAIN 88 96 GLY-RICH (HINGE REGION) .
CC FT DOMAIN 119 192 RNA-BINDING (RRM) 2.
CC FT DOMAIN 206 356 ARG/SER-RICH (RS DOMAIN).
CC FT VARSPIC 318 338 Missing (in isoform short).
CC FT VARSPIC 318 338 /FTId=VSP_005878.
CC FT CONFLICT 74 74 T->S (IN REF. 1).
CC FT CONFLICT 102 106 MISSING (IN REF. 1).
CC FT CONFLICT 195 195 A->R (IN REF. 1).
CC FT CONFLICT 228 228 T->S (IN REF. 1).
CC FT CONFLICT 260 260 A->R (IN REF. 1).
CC FT CONFLICT 279 281 APV->RSR (IN REF. 1).
CC FT CONFLICT 293 293 S->T (IN REF. 1).
CC SQ SEQUENCE 375 AA; 42393 MW; 20BA327CB4A4194E CRC64;

```


Query Match 45.7%; Score 42; DB 1; Length 375;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 STRGRSRGXSGRSQS 19
Db 202 SGRGRSRSSSSRSRS 216

Search completed: September 28, 2004, 06:05:48
Job time : 8.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-4

Perfect score: 92

Sequence: 1 SHQESTRGSRGSRGSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SPTEMBL_25:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virulus:*
17: sp_bacteriaph:*
18: sp_archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	97.8	797	4	Q16824
2	90	97.8	990	4	Q15206
3	90	97.8	1218	4	Q05331
4	84	91.3	798	4	Q9H4U3
5	84	91.3	1084	4	Q01212
6	75	81.5	591	4	Q01720
7	75	81.5	687	4	Q9H4U2
8	69	75.0	322	4	Q03838
9	69	75.0	800	5	Q15845
10	50	54.3	822	3	Q9P312
11	48	52.2	1229	3	Q86ZHI
12	48	52.2	260	10	Q9FMU3
13	47	51.1	260	10	Q9FMU3
14	47	51.1	260	10	Q7XD54
15	47	51.1	471	5	Q9UAY0
16	47	51.1	619	13	Q7J397

17	47	51.1	983	11	Q8B5T3	Q8b5t3 mus musculu
18	47	51.1	983	11	Q812D4	Q812d4 mus musculu
19	47	51.1	1014	11	Q8B1T6	Q8b1t6 mus musculu
20	47	51.1	1015	11	Q8B1W4	Q8b1w4 mus musculu
21	47	51.1	1035	11	Q812D5	Q812d5 mus musculu
22	47	51.1	1067	11	Q8CIG7	Q8cig7 mus musculu
23	47	51.1	1105	11	Q8X102	Q8x102 mus musculu
24	47	51.1	1107	11	Q80X50	Q80x50 mus musculu
25	47	51.1	1112	11	Q8B0U1	Q8b0u1 mus musculu
26	47	51.1	1284	5	Q9V9Y3	Q9v9y3 drosophila
27	46	50.0	373	10	Q23876	Q23876 oryza sativ
28	46	50.0	688	13	Q42375	Q42375 brachydanto
29	46	50.0	836	10	Q9SKZ9	Q9skz9 arabidopsis
30	46	50.0	1087	4	Q9BTU3	Q9btu3 homo sapien
31	46	50.0	1829	16	Q86808	Q86808 streptomyce
32	45.5	49.5	144	5	Q8T9K6	Q8t9k6 drosophila
33	45.5	49.5	433	5	Q26649	Q26649 strongyloce
34	45	48.9	198	16	Q8VKK4	Q8vkk4 mycobacteri
35	45	48.9	476	12	Q80890	Q80890 herpesviru
36	45	48.9	652	5	Q77168	Q77168 apis mellif
37	45	48.9	1142	13	Q7SZV2	Q7szv2 xenopus lae
38	45	48.9	1394	10	Q8H8L9	Q8h8l9 oryza sativ
39	45	48.9	1433	10	Q94ZG3	Q94zg3 oryza sativ
40	44.5	48.4	319	10	Q94A20	Q94a20 arabidopsis
41	44.5	48.4	553	5	Q9U203	Q9u203 caenorhabdi
42	44.5	48.4	659	5	Q9U202	Q9u202 caenorhabdi
43	44	47.8	157	16	Q7VFP1	Q7vfp1 helicobacte
44	44	47.8	188	10	Q8H8P6	Q8h8p6 oryza sativ
45	44	47.8	222	4	Q96B40	Q96b40 homo sapien

ALIGNMENTS

Q16824	PRELIMINARY;	PRT;	797 AA.
AC Q16824;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE Profilaggrin (Fragment).			
GN FLG.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=91064347; PubMed=2248957;			
FA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.,			
RT "Organization, structure, and polymorphisms of the human profilaggrin			
RT gene [published erratum appears in Biochemistry 1991 Jun			
RT 11;30(23):5814.";			
RT 11;30(23):5814.";			
RL Biochemistry 29:9432-9440 (1990).			
DR EMBL; M60502; AAA63248.1;			
DR GO; GO:0005196; F:structural molecule activity; IRA.			
DR InterPro; IPR003303; Flaggrin.			
DR Pfam; PF03516; Flaggrin; 4.			
DR PRINTS; PR00487; FLAGGRIN.			
FT NON TER			
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDA858 CRC64;			
Qy	1	SHQESTRGSRGSRGSGS 19	97.8%; Score 90; DB 4; Length 797;
Db	427	SHQESTRGSRGSRGSGS 445	Best Local Similarity 94.7%; Pred. No. 1.6e-06;
			Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RESULT 2			

015206 ID Q15206 PRELIMINARY; PRT; 990 AA.
 AC Q15206;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9106347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]."
 RT 11.30(23):5814;
 RL Biochemistry 29:9432-9440(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9125519; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene."
 RL Biochemistry 30:5814-5814(1991).
 DR EMBL; M60494; AA63244.1;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON TER 990
 SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
 Query Match 97.8%; Score 90; DB 4; Length 990;
 Best Local Similarity 94.7%; Pred. No. 2.1e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SH0ESTRGSRGSGSGS 19
 DB 227 SH0ESTRGSRGSGSGS 245
 RESULT 3
 005331 ID Q05331 PRELIMINARY; PRT; 1218 AA.
 AC Q05331;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE FILAGGRIN (PROFILAGGRIN) (Fragment).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FORESKIN;
 RX MEDLINE=93109348; PubMed=8417156;
 RA Markova N.G., Marekov L.N., Chipkev C.C., Gan S.-Q., Idler W.W.,
 Steinert P.M.;
 RT "Profilaggrin is a major epidermal calcium-binding protein."
 RL Mol. Cell. Biol. 13:613-625(1993).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 CC EMBL; M96943; AA36487.1; -
 CC FIR; A48118; A48118.
 CC HSSP; P02593; 1CDM.
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0005509; F:calcium ion binding; TAS.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 6.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100 CABP; 1.
 KM Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
 KM Polymorphism.
 FT CA_BIND 19 32 SITE I (BY SIMILARITY).
 FT CA_BIND 62 73 SITE II (BY SIMILARITY).
 FT NON TER 1218 1218
 SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;
 Query Match 97.8%; Score 90; DB 4; Length 1218;
 Best Local Similarity 94.7%; Pred. No. 2.6e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SH0ESTRGSRGSGSGS 19
 DB 449 SH0ESTRGSRGSGSGS 467
 RESULT 4
 09H4U3 ID Q9H4U3 PRELIMINARY; PRT; 798 AA.
 AC Q9H4U3;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL356504; CAC13171.1; -
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 4.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON TER 1
 SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
 Query Match 91.3%; Score 84; DB 4; Length 798;
 Best Local Similarity 89.5%; Pred. No. 1.7e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SH0ESTRGSRGSGSGS 19
 DB 428 SH0ESTRGSRGSGSGS 446
 RESULT 5
 001212 ID Q01212 PRELIMINARY; PRT; 1084 AA.

```

AC Q01212; Q03840;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., Koberide O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published extratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814.];"
RL Biochemistry 23:9432-9440(1990).
DR EMBL: M60503; AAA63243.1; JOINED.
DR EMBL: M60501; AAA63243.1; JOINED.
DR GO: GO:0005882; C:Intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007275; F:development; NAS.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF03516; Filaggrin; 6.
DR PRINTS: PR00487; Filaggrin.
DR NON_TER 1
FT SEQUENCE 1084 AA; 115271 MW; 80C64608BD5A362D CRC64;
SQ

Query Match
Best Local Similarity 91.3%; Score 84; DB 4; Length 1084;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
Db 64 SHOESTRGRSGRSGSGS 82
|||||
ID Q01720 PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Ntunusukisiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
RL J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTOOLITICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL: L01089; AAA60177.1; -.
DR EMBL: L01090; AAA60176.1; -.

```

```

DR PIR: A45135; A45135.
DR PIR: A48118; A48118.
DR HSSP: P80511; 1E8A.
DR MIM: 135940; -.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001751; CABP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Filaggrin; 2.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; Filaggrin.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467 POTENTIAL.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FILAGGRIN.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 591
FT SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;
SQ

Query Match
Best Local Similarity 81.5%; Score 75; DB 4; Length 591;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
Db 449 SHOESTRGRSGRSGSGS 467
|||||
ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL: AL356504; CAC13172.1; -.
DR PIR: A48118; A48118.
DR HSSP: P80511; 1E8A.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001751; CABP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Filaggrin; 3.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; FILAGGRIN.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
FT NON_TER 687
FT SEQUENCE 687 AA; 76659 MW; 8000363FBBF07B74 CRC64;
SQ

Query Match
Best Local Similarity 81.5%; Score 75; DB 4; Length 687;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 SHOESTRGRSGSGSGS 19
 DB 449 SHOESTRGRSGSGSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT: 465 AA.

AC 003838.
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE FILAGRIN (PROFILAGRIN) (Fragment).
 GN FLG.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;
 MEDLINE=91064347; PubMed=2248957;

RA Can.S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 29:9432-9440(1990).
 [2]

RP REVISIONS.

RX MEDLINE=91255199; PubMed=2043621;
 RA Can.S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
 PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
 FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND
 WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
 REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
 INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
 COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN
 GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 PROTEOLYTICALLY CLEAVED.
 DR EMBL; M60499; AAA63246.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 3.
 DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1
 FT SEQUENCE 465 AA; 50280 MW; C683744C5E134097 CRC64;

Query Match 76.1%; Score 70; DB 4; Length 465;
 Best Local Similarity 78.9%; Pred. No. 0.0021;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGSGSGS 19
 DB 227 SHOESTRGRSGSGSGS 245

RESULT 9
 ID 075370 PRELIMINARY; PRT: 322 AA.

AC 075370.
 DT 01-NOV-1998 (TRENBLREL. 08, Created)

DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Epidermal filaggrin (Fragment).
 GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RX MEDLINE=99101527; PubMed=9886436;

RA Girbal-Neuhausser E., Durieux J.J., Arnaud M., Dalbon P., Sebtag M.,
 RA Vincent C., Simon M., Sersu T., Masson-Bessiere C.,
 RA Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated
 anti-filaggrin autoantibodies are posttranslationally generated on
 various sites of (pro)filaggrin by deamination of arginine residues."
 RL J. Immunol. 162:585-594(1999).
 DR EMBL; AF04380; AAC2355.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 2.
 DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1
 FT SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match 75.0%; Score 69; DB 4; Length 322;
 Best Local Similarity 77.8%; Pred. No. 0.002;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGSGSGS 18
 DB 305 SHOESTRGRSGSGSGS 322

RESULT 10
 ID 015845 PRELIMINARY; PRT: 800 AA.

AC 015845.
 DT 01-JAN-1998 (TRENBLREL. 05, Created)

DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE IL439.4.

GN IL439.4.
 OS Leishmania major.

OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_TaxID=5664;

RP SEQUENCE FROM N.A.
 RC STRAIN=Friledin;

RX MEDLINE=99179887; PubMed=10077609;
 RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,

RA Wagness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
 RA Bastien P., Fu G., Ivans A., Stuart K.;

RT "Leishmania major Friledin chromosome 1 has an unusual distribution of
 protein-coding genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).

DR EMBL; AE001274; AAC24675.1; -;
 DR FIR; 102852; 102852.
 DR SEQUENCE 800 AA; 83782 MW; DA57AE627913CBAC CRC64;

Query Match 54.3%; Score 50; DB 5; Length 800;
 Best Local Similarity 55.6%; Pred. No. 8.4;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 HOESTRGRSGSGSGS 19
 DB 429 HRDVGRLSTASGSGS 446

RESULT 11

ID 09P312 PRELIMINARY; PRT: 822 AA.

AC 09P312.
 DT 01-OCT-2000 (TRENBLREL. 15, Created)

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.

GN B12P1.10.

```

OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -.
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR005120; Smg4_UBF3.
DR Pfam; PF00076; rrm1_1.
DR Pfam; PF03467; Smg4_UBF3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 52.2%; Score 48; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTGRSGXSGRSGS 19
Db 414 RESASGRTRRRRCGR 430

RESULT 12
ID 086ZHL PRELIMINARY; PRT; 1229 AA.
AC 086ZHL;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Related to cell division cycle 2-related protein kinase 7.
GN 754.060.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX294020; CAD70910.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0000910; F:cyclin kinase; IEA.
DR InterPro; IPR000719; Prc1_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00060; Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

```

```

KW Cell division; Kinase.
SQ SEQUENCE 1229 AA; 136776 MW; 43BE612656F73E4 CRC64;

Query Match 52.2%; Score 48; DB 3; Length 1229;
Best Local Similarity 57.9%; Pred. No. 29;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTGRSGXSGRSGS 19
Db 242 SHHERRSRDRKRSRSGRS 260

RESULT 13
ID Q9FWU3 PRELIMINARY; PRT; 260 AA.
AC Q9FWU3;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative secretory protein.
GN OSUNBA0051D19.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RC Buehl C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsieh J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUNBA0051D19 genomic sequence.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC023240; AAG33529.1; -.
DR Gramene; Q9FWU3; -.
DR InterPro; IPR007541; BSP.
DR Pfam; PF04450; BSP; 1.
SQ SEQUENCE 260 AA; 28706 MW; CDC696A4EC58B022 CRC64;

Query Match 51.1%; Score 47; DB 10; Length 260;
Best Local Similarity 69.2%; Pred. No. 7.9;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRGRSRSGXSGRSG 18
Db 35 SKGRSRGAHGRRG 47

RESULT 14
ID Q7XD54 PRELIMINARY; PRT; 260 AA.
AC Q7XD54;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative secretory protein.
GN OSUNBA0051D19.12.
OS Oryza sativa (japanese cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RC The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Ming R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017108; AAP54393.1; -
 SQ SEQUENCE 260 AA; 28706 MW; CDC696A4EC58B022 CRC64;

Query Match 51.1%; Score 47; DB 10; Length 260;
 Best Local Similarity 69.2%; Pred. No. 7.9;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRGSRGSGSGSG 18
 :|||||
 Db 35 SRGSRGAGRRG 47

RESULT 15

Q9UAY0 PRELIMINARY; PRT; 471 AA.
 AC Q9UAY0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE W03G1.5 protein.
 GN W03G1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copestake T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten N., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons V., Percy C., Riiken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A., Scheet P., Harper M.;
 RT "The sequence of C. elegans cosmid W03G1."
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF125964; AAD14753.1; -
 DR PIR; T33997; T33997
 DR WormPep; W03G1.5; CE17283.
 SQ SEQUENCE 471 AA; 50885 MW; BDF30B59464A985B CRC64;

Query Match 51.1%; Score 47; DB 5; Length 471;
 Best Local Similarity 52.6%; Pred. No. 15;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTRGSRGSGSGSGS 19
 :|||||
 Db 167 SSRSPSRGRRGSGSGSGS 185

Search completed: September 28, 2004, 06:12:42
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-4
Perfect score: 92
Sequence: 1 SHQESTRGRSGXSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_29Jan04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
9: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	2	AAW61508 Peptide c
2	90	97.8	19	2	AAW61514 Peptide c
3	90	97.8	1467	5	ABB97605 Novel hum
4	85	93.5	19	2	AAW61517 Peptide c
5	85	92.4	19	2	AAW61516 Peptide c
6	84	91.3	19	2	AAW61506 Peptide c
7	84	91.3	19	2	AAW61515 Peptide c
8	84	91.3	19	2	AAW61507 Peptide c
9	84	91.3	19	2	AAW61512 Peptide c
10	84	91.3	19	2	AAW61509 Peptide c
11	84	91.3	19	2	AAW61505 Peptide c
12	78	84.8	19	2	AAW61511 Peptide c
13	78	84.8	19	2	AAW61513 Peptide c
14	78	84.8	19	2	AAW61510 Peptide c
15	75	81.5	21	2	AAW61520 Peptide c
16	75	81.5	22	4	AAE07235 IGP1546 P
17	66	71.7	330	2	AAV22956 Human fil
18	66	71.7	330	2	AAV22955 Human fil
19	66	71.7	330	2	AAV22957 Human fil
20	56	67.4	330	2	AAV22954 Human fil
21	56	60.9	477	6	ABO07142 Novel hum
22	51	55.4	1711	4	AAW79819 Human pro
23	51	55.4	1951	4	AAW78835 Human pro
24	49	53.3	641	4	ABG19110 Novel hum
25	48	52.2	772	4	AA664573 Human Mre

26	47	51.1	184	3	AAW57041 Human pro
27	47	51.1	1199	4	ABB58274 Drosophila
28	46	50.0	134	4	ABG23065 Novel hum
29	46	50.0	318	6	ABU70492 Human adi
30	46	50.0	903	6	ABU70701 Human adi
31	45.5	49.5	135	4	ABB66181 Drosophila
32	44	47.8	124	4	ABB72046 Drosophila
33	44	47.8	222	4	AAW93652 Human pol
34	44	47.8	229	7	ADE59825 Rat Prote
35	44	47.8	229	7	ADE59833 Rat Prote
36	44	47.8	229	7	ADE59837 Rat Prote
37	44	47.8	229	7	ADE59829 Rat Prote
38	44	47.8	285	4	ABB69359 Drosophila
39	44	47.8	826	3	AAV96513 Human zsi
40	44	47.8	837	4	ABG19458 Novel hum
41	44	47.8	1038	7	ADC03412 Rice flow
42	44	47.8	1308	6	ABR57072 Human tau
43	44	47.8	1331	7	ADC10010 Human NOV
44	43	46.7	173	6	ABR41796 Human DIT
45	43	46.7	356	5	AAE15535 Beta vulg

ALIGNMENTS

RESULT 1	AAW61508	standard, peptide, 19 AA.
ID	AAW61508	
XX	AAW61508;	
AC	AAW61508;	
DT	26-OCT-1998	(first entry)
XX		
DE	Peptide cfc4, based on cDNA of a profilaggrin repeat.	
XX		
KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
KW	solid phase synthesis; peptide amide; polyclonal antibody;	
KW	monoclonal antibody.	
XX		
OS	Synthetic.	
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Modified-site	13
XX		/note= "Citruiline"
PN	WO9822503-A2.	
XX		
PD	28-MAY-1998.	
XX		
PF	14-NOV-1997;	97WO-NL000624.
XX		
PR	15-NOV-1996;	96NL-01004539.
XX		
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
XX	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
PI	Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;	
XX		
DR	WPI, 1998-398613/34.	
XX		
PT	Peptide derived from an antigen recognised by autoantibodies - is	
PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
PT	used in diagnosis of the disease.	
XX		
PS	Disclosure; Page 6; 19pp; English.	
XX		
CC	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of	
CC	the profilaggrin antigen which is recognised by autoantibodies from	
CC	patients with Rheumatoid arthritis (RA). This peptide is reactive with a	
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	The peptides were created by using standard solid phase synthesis, which	
CC	produced them as peptide amides. These sequences may be used in the	

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 1 SHQESTRGRSGXSGRSGS 19

RESULT 2

AAW61514
 ID AAW61514 standard; peptide; 19 AA.

AC AAW61514;

DT 26-OCT-1998 (first entry)

DE Peptide cf, based on cDNA of a proflilagrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.1e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 1 SHQESTRGRSGXSGRSGS 19

RESULT 3

ABB97605
 ID ABB97605 standard; protein; 1467 AA.

XX ABB97605;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

KW Human; anti-nausea; vulnary; anti-inflammatory; immunomodulator;
 KW anti-fertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

XX Homo sapiens.

XX W0200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RR;

XX WPI; 2002-292408/33.

XX N-PSDB; ABB97605.

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX
 SQ Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.5e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 449 SHQESTRGRSGXSGRSGS 1467

AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

DT 26-OCT-1998 (first entry)

DE Peptide cf0, based on cDNA of a proflilagrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

XX

```

OS Synthetic.
OS Homo sapiens.
XX WO9822503-A2.
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX PS Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX SQ Sequence 19 AA;
XX
XX Query Match 93.5%; Score 86; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. NO. 9.1e-07;
XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHOESTGRSRGSGRSGS 19
DB 1 SHOESTGRSRGSGRSGS 19

```

```

XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX PS Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX SQ Sequence 19 AA;
XX
XX Query Match 92.4%; Score 85; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. NO. 1.3e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 SHOESTGRSRGSGRSGS 19
DB 1 SHOESTGRSRGSGRSGS 19

```

```

RESULT 5
AAW61516
ID AAW61516 standard; peptide; 19 AA.
XX
XX AAW61516;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cFE, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX PN WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX

```

```

RESULT 6
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX
XX AAW61506;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cFE2, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 9
XX FT /note= "Ctrlulline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX
XX PS Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 7
 AAW61515
 ID AAW61515 standard; peptide; 19 AA.

AC AAW61515;

DT 26-OCT-1998 (first entry)

DE Peptide cFA, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KW monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

FN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

Db 1 SHOESTRGRSGRSGSGS 19
 RESULT 8
 ID AAW61507
 ID AAW61507 standard; peptide; 19 AA.

AC AAW61507;

DT 26-OCT-1998 (first entry)

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KW monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

FN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 9
 AAW61512

ID AAW61512 standard; peptide; 19 AA.

AC AAW61512;

DT 26-OCT-1998 (first entry)

DE Peptide cfc8, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7
 FT Modified-site /note= "Citruilline"
 FT Modified-site 13
 FT /note= "Citruilline"
 PN WO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;
 Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 1.9e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHOESTRGSRGSGSGS 19
 DB 1 SHOESTRGSRGSGSGS 19
 RESULT 10
 AAM61509
 ID AAM61509 standard; peptide; 19 AA.
 AC AAM61509;
 XX
 XX 26-OCT-1998 (first entry)
 DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 16
 FT Modified-site /note= "Citruilline"
 PN WO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;
 Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHOESTRGSRGSGSGS 19
 DB 1 SHOESTRGSRGSGSGS 19
 RESULT 11
 AAM61505
 ID AAM61505 standard; peptide; 19 AA.
 AC AAM61505;
 XX
 XX 26-OCT-1998 (first entry)
 DE Peptide cfc1, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7
 FT Modified-site /note= "Citruilline"
 PN WO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.
 XX

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PS used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 91.3%; Score 84; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.9e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHQESTGRGRSGRSGSGS 19
DB 1 SHQESTGRGRSGRSGSGS 19
XX
XX RESULT 12
XX AAW61511
XX ID AAW61511 standard; peptide; 19 AA.
XX
XX AAW61511;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
FT
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 84.8%; Score 78; DB 2; Length 19;
XX Best Local Similarity 84.2%; Pred. No. 1.7e-05;
XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SHQESTGRGRSGRSGSGS 19
DB 1 SHQESTGRGRSGRSGSGS 19
XX
XX RESULT 13
XX AAW61513
XX ID AAW61513 standard; peptide; 19 AA.
XX
XX AAW61513;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc9, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 16 /note= "Citruilline"
FT
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC

```

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
 |||||
 Db 1 SHOESTXGRSGRSGSGS 19

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 DT 26-OCT-1998 (first entry)
 DE Peptide cfc6, based on cDNA of a proflilagrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97MO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDTG ONDERZOEK IN NEDER.
 XX (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.7e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 SHOESTGRSGRSGSGS 19
 |||||
 Db 1 SHOESTXGRSGRSGSGS 19

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 DT 26-OCT-1998 (first entry)
 DE Peptide XI based on cDNA of a proflilagrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97MO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDTG ONDERZOEK IN NEDER.
 XX (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 5.6e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HOESTGRSGRSGSGS 19
 |||||
 Db 4 HOESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:13
 Job time : 55.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds

(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-4

Perfect score: 92

Sequence: 1 SHQSTRGRSGXSGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-4
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-3
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-8
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-7
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	56	60.9	477	15	US-10-161-927-62	Sequence 62, Appl
17	51	55.4	78	16	US-10-437-963-203073	Sequence 203073,
18	50	54.3	133	16	US-10-437-963-136367	Sequence 136367,
19	49	53.3	889	16	US-10-437-963-181020	Sequence 181020,
20	48	52.2	772	15	US-10-369-493-6474	Sequence 6474, Ap
21	47	51.1	133	16	US-10-767-701-3436	Sequence 3436, A
22	47	51.1	166	16	US-10-437-963-200529	Sequence 200529,
23	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
24	47	51.1	258	16	US-10-437-963-140633	Sequence 140633,
25	47	51.1	952	16	US-10-437-963-173648	Sequence 173648,
26	46.5	50.5	324	12	US-10-425-114-48469	Sequence 48469, A
27	46.5	50.5	332	12	US-10-425-114-66862	Sequence 66862, A
28	46	50.0	373	16	US-10-437-963-143403	Sequence 143403,
29	46	50.0	466	15	US-10-369-493-8178	Sequence 8178, Ap
30	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
31	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
32	45	48.9	49	12	US-10-424-599-187406	Sequence 187406,
33	45	48.9	105	16	US-10-437-963-196208	Sequence 196208,
34	45	48.9	128	16	US-10-437-963-111390	Sequence 111390,
35	45	48.9	162	16	US-10-437-963-137223	Sequence 137223,
36	45	48.9	248	16	US-10-437-963-185716	Sequence 185716,
37	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
38	45	48.9	1232	16	US-10-437-963-192222	Sequence 192222,
39	45	48.9	1332	15	US-10-374-780A-1037	Sequence 1037, Ap
40	45	48.9	1360	16	US-10-437-963-185720	Sequence 185720,
41	45	48.9	1394	16	US-10-437-963-185722	Sequence 185722,
42	45	48.9	1409	15	US-10-374-780A-1038	Sequence 1038, Ap
43	45	48.9	1409	16	US-10-437-963-110664	Sequence 110664,
44	45	48.9	1433	15	US-10-374-780A-1035	Sequence 1035, Ap
45	45	48.9	1433	15	US-10-374-780A-1040	Sequence 1040, Ap

ALIGNMENTS

RESULT 1
US-09-308-150-4 : Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
: APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
: APPLICANT: Schellekens, Gerardus Antonius
: APPLICANT: Raats, Jozef Maria Hendrik
: APPLICANT: Hoel, Rene Michael Antonius
: APPLICANT: Stichting Scheikundig Onderzoek Nederland
: TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
: TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
: FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
: CURRENT APPLICATION NUMBER: US/09/308,150
: PRIOR FILING DATE: 1999-09-30
: PRIOR APPLICATION NUMBER: PCT/NL97/00624
: PRIOR FILING DATE: 1997-11-14
: PRIOR APPLICATION NUMBER: NL 1004539
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 19
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:Derived from
: OTHER INFORMATION: known cDNA sequences of human profilaggrin
: OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4
Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 2

US-09-308-150-11
; Sequence 11, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2,6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 3

US-09-308-150-14
; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 9,9e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 4

US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1,4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 5

US-09-308-150-1
; Sequence 1, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 6
US-09-308-150-2
/ Sequence 2, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 7
US-09-308-150-3
```

```
/ Sequence 3, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 8
US-09-308-150-5
/ Sequence 5, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentn Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```

US-09-308-150-5
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 9
US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-8
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 1.9e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 10
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

US-09-308-150-6
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-6
Query Match          84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 12
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSGRSGS 19
Db 1 SHOESTRGSRGSGRSGS 19

RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Ventrcooij, Walthertus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: AUTOANTIBODIES DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSGRSGS 19
Db 1 SHOESTRGSRGSGRSGS 19

RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Ventrcooij, Walthertus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
NAME/KEY: DISULFID
LOCATION: (3)..(16)
US-09-308-150-10

Query Match 81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00043;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HOESTRGSRGSGRSGS 19
Db 4 HOESTRGSRGSGRSGS 21

RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydie
TITLE OF INVENTION: PEPTIDS DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.003INPUS00 INNS:031
CURRENT APPLICATION NUMBER: US/09/747,029A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQESTRGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR78:*
2: PIR7:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	93.5	2248	2 A35938	profilaggrin - hum
2	84	91.3	416	2 A32947	filaggrin precursor
3	75	81.5	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2ML47	E2 protein - human
5	48	52.2	772	2 T27512	hypothetical prote
6	48	52.2	822	2 T51049	related to nucleol
7	47	51.1	308	2 A29379	glycine-rich prote
8	47	51.1	373	2 T02976	probable DNA bindi
9	47	51.1	471	2 T33997	hypothetical prote
10	46	50.0	467	2 T01462	hypothetical prote
11	46	50.0	759	2 F86362	filaggrin-like prote
12	46	50.0	836	2 G84727	probable DNA topoi
13	45	48.9	409	2 T35118	probable secreted
14	45	48.9	1703	2 T15047	SNF2 protein - yea
15	44.5	48.4	682	2 T15092	hypothetical prote
16	44.5	48.4	687	1 S13654	ATP-dependent RNA
17	44.5	48.4	1496	2 T00499	probable reticulom
18	44	47.8	229	2 J07219	nuclear protein SR
19	44	47.8	452	2 S36482	E2 protein - human
20	44	47.8	483	2 S36470	E2 protein - human
21	44	47.8	485	2 A40751	finger protein MZF
22	44	47.8	867	2 T27136	hypothetical prote
23	44	47.8	871	2 T27135	hypothetical prote
24	44	47.8	1232	2 S40766	hypothetical prote
25	43.5	47.3	463	2 T51194	hypothetical prote
26	43	46.7	123	2 T16234	hypothetical prote
27	43	46.7	151	2 T21827	hypothetical prote
28	43	46.7	210	2 C84404	hypothetical prote
29	43	46.7	237	2 T13649	hypothetical prote

30	43	46.7	240	2 A46179	U2 snRNP auxiliary
31	43	46.7	488	2 S36570	E2 protein - human
32	43	46.7	514	1 W2ML5	E2 protein - human
33	43	46.7	695	2 T51652	dsRNA-binding prot
34	43	46.7	718	2 T02531	probable SCARECROW
35	43	46.7	745	2 G01025	serine/threonine p
36	43	46.7	774	2 T48609	probable serine/th
37	43	46.7	776	2 A87330	sensory box histid
38	43	46.7	800	2 T02852	probable membrane
39	43	46.7	800	2 T51653	dsRNA-binding prot
40	43	46.7	843	2 H86209	protein F2G5.10 [
41	43	46.7	1119	2 T16720	hypothetical prote
42	43	46.7	1339	2 T47841	hypothetical prote
43	43	46.7	2233	2 S63347	acetyl-CoA carboxy
44	42	45.7	166	2 T39586	rna binding protei
45	42	45.7	184	2 S26046	sex-determining pr

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gam, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; MUID:91064347; PMID:2248957
A/Accession: A35938
A/Status: Preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:J02929
C/Genetics:
A/Genes: GDB:FIG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: filaggrin repeat
F:570-893/Region: filaggrin repeat
F:1074-1197/Region: filaggrin repeat
F:1573-1896/Region: filaggrin repeat

Query Match 93.5%; Score 86; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred.No. 4.3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSRGRSGXSGS 19
DB 551 SHQESTRGRSRGRSGS 569

RESULT 2
A32947
filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinnley-Grant, L.U.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of t
A/Reference number: A32947; MUID:89296601; PMID:2743331
A/Accession: A32947
A/Status: Preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCR>
A/Cross-references: GB:M2435; NID:9182604; PIDN:AA52454.1; PID:9182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

```
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
A:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match
  Best Local Similarity 91.3%; Score 84; DB 2; Length 416;
  Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 SHOESTRGRSGRSGSGS 19
    |||||
Db 7 SHOESTRGRSGRSGSGS 25

RESULT 3
A45135
Profilaggrin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A45135
R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nurusksirtl, W.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992
A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A:Reference number: A45135; MUID:93054736; PMID:1429717
A:Accession: A45135
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-591 <PRE>
A:Cross-references: GB:101089; GB:M90967; NID:9190408; PIDN:AAA60177.1; PID:553621
A:Note: sequence extracted from NCBI backbone (NCBIF:118773)
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match
  Best Local Similarity 81.5%; Score 75; DB 2; Length 591;
  Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 SHOESTRGRSGRSGSGS 19
    |||||
Db 449 SHOESTRGRSGRSGSGS 467

RESULT 4
M2ML47
E2 protein - human papillomavirus type 47
C:Species: human papillomavirus type 47
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: D35324
R:Kiyono, T.; Adachi, A.; Ishibashi, M.
Virology 177, 401-405, 1990
A:Title: Genome organization and taxonomic position of human papillomavirus type 47 inf
A:Reference number: A35324; MUID:90281611; PMID:2162112
A:Accession: D35324
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-506 <KTY>
A:Cross-references: GB:M42305; NID:9333062; PIDN:AAA46979.1; PID:9333067
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match
  Best Local Similarity 63.0%; Score 58; DB 1; Length 506;
  Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy 1 SHOESTRGRSGRSGSGS 19
    |||||
Db 342 SRENTGRGRGRGRGAGS 360

RESULT 5
T27512
Hypothetical protein ZC302.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2003
C:Accession: T27512
R:Kelly, P.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z20380
A:Accession: T27512
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-772 <ML>
A:Cross-references: EMBL:T273978; PIDN:CAA98292.1; GSPDB:GN00023; CESP:ZC302.1
A:Experimental source: clone ZC302
C:Genetics:
A:Gene: CESP:ZC302.1
A:Map position: 5
A:introns: 43/2; 70/3; 94/2; 129/2; 393/3; 548/3; 598/3; 642/3
C:Superfamily: double-strand break repair protein MRE11; phosphodiesterase core homology

Query Match
  Best Local Similarity 52.2%; Score 48; DB 2; Length 772;
  Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Cy 1 SHOESTRGRSGRSGSGS 19
    |||||
Db 697 SKQPTTRGRGRGRGAGS 715

RESULT 6
T51049
related to nucleolar phosphoprotein [imported] - Neurospora crassa
N:Alternate names: protein B12Fl.10
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51049
R:Schulte, U.; Aign, V.; Hohenseil, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51049
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-822 <SCH>
A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12Fl.10
A:Experimental source: BAC clone B12Fl; strain OR74A
C:Genetics:
A:Gene: NCSP:B12Fl.10
A:Map position: 6
A:introns: 80/2

Query Match
  Best Local Similarity 52.2%; Score 48; DB 2; Length 822;
  Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy 3 OESTRGRSGRSGSGS 19
    |||||
Db 414 RBSAGRTGRGRGRGT 430

RESULT 7
A29379
glycine-rich protein GRP3 - brine shrimp
C:Species: Artemia salina (brine shrimp)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Feb-1997
C:Accession: A29379
R:Cruz-Alvarez, M.; Pellicer, A.
J. Biol. Chem. 262, 13377-13380, 1987
A:Title: Cloning of a full-length complementary DNA for an Artemia salina glycine-rich p
A:Reference number: A29379; MUID:88007550; PMID:2443491
```


A:Accession: A29379
A:Molecule type: mRNA
A:Residues: 1-308 <CRU>
C:Keywords: nucleus

Query Match 51.1%; Score 47; DB 2; Length 308;
Best Local Similarity 75.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 RGRSGRSGSGS 18
Db 210 RGRSGRSGSGS 221

RESULT 8

T02976

probable DNA binding protein PCF2 - rice

C:Species: Oryza sativa (rice)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000

C:Accession: T02976

R:Kosugi, S.; Ohashi, Y.

Plant Cell 9, 1607-1619, 1997

A>Title: PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell

A:Reference number: Z14803; MUID:97480096; PMID:9338963

A:Accession: T02976

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-373 <ROS>

A:Cross-references: EMBL:D87261; NID:G2580439; PIDN:BAA21143.1; PID:G2580440

A:Experimental source: cultivar Nipponbare

Query Match 51.1%; Score 47; DB 2; Length 373;
Best Local Similarity 53.3%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGS 15
Db 334 SHOESTRGRSGRSGS 348

RESULT 9

T33997

hypothetical protein W03G1.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33997

R:Pauley, A.; Schneet, P.; Harper, M.

submitted to the EMBL Data Library, February 1999

A>Description: The sequence of C. elegans cosmid W03G1.

A:Reference number: Z21454

A:Accession: T33997

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-471 <PAU>

A:Cross-references: EMBL:AF129664; PIDN:AAD14753.1; GSPDB:GN00022; CESP:W03G1.5

A:Experimental source: strain Bristol N2; clone W03G1

C:Genetics:

A:Gene: CESP:W03G1.5

A:Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGS 19
Db 167 SHRSPGRSGRSGS 185

RESULT 10

T01462

hypothetical protein T24H24.19 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Mar-1999

C:Accession: T01462

R:Courteney, L.; Stoneking, T.; Langston, Y.; Mead, K.

submitted to the EMBL Data Library, August 1998

A>Description: The sequence of A. thaliana T24H24.

A:Reference number: Z14333

A:Accession: T01462

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-467 <COU>

A:Cross-references: EMBL:AF075598; NID:G3293581; PID:G3377836

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

Query Match 50.0%; Score 46; DB 2; Length 467;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 2 HOESTRGRSGRSGS 19
Db 188 HNNRGRGRSGRSGS 205

RESULT 11

F6362

F19G10.17 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001

C:Accession: F6362

R:Rheologis, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, C.M.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, X.; Liu, Z.A.; Luros, J.S.; Matli, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A66141; MUID:21016719; PMID:11130712

A:Accession: F6362

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-759 <STO>

A:Cross-references: GB:AE005172; NID:G2462837; PIDN:ABE72172.1; GSPDB:GN00141

A:Map position: 1

Query Match 50.0%; Score 46; DB 2; Length 759;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 RGRSGRSGS 17
Db 728 RGRSGRSGS 738

RESULT 12

G84727

probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84727

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

Wu, X.; Xie, H.; Moffet, K.S.; Cronin, L.A.; Shen, M.; Vankar, S.E.; Umayam, L.; Tallon, I.

euus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-766, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A64420; MUID:20083487; PMID:10617197

A:Accession: G84727

A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-836 <STD>
A:Cross-references: GB:AE020293; NID:g4263718; PIDN:AAD15404.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT293200
A:Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 RGRSRGRSGXSGS 19
Db 812 RGRSRGRSGXSGS 824

RESULT 13

135118
Probable secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 08-Sep-2000
C:Accession: T15118
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, March 1998
A:Reference number: 221568
A:Accession: T15118
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-409 <SEP>
A:Cross-references: EMBL:AL022268; PIDN:CAA18335.1; GSPDB:GN00070; SCOEDB:SC4H2.19C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC4H2.19C
C:Superfamily: Streptomyces coelicolor probable secreted protein SC4H2.19C

Query Match 48.9%; Score 45; DB 2; Length 409;
Best Local Similarity 60.0%; Pred. No. 18;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 5 STRGRSRGRSGXSGS 19
Db 9 STRGRSRGRSGXSGS 23

RESULT 14

515047
SNF2 protein - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5509; protein O5509c; protein YOR290C
C:Species: Saccharomyces cerevisiae
C>Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Sep-1999
C:Accession: S15047; S16820; S67192; S67194; S72058
R:Laurent, B.C.; Treitel, M.A.; Carlson, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 2687-2691, 1991
A:Title: Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in trans
A:Reference number: S15047; MUID:91187857; PMID:1901413
A:Accession: S15047
A:Molecule type: DNA
A:Residues: 1-1703 <PRO>
A:Cross-references: EMBL:M61703; NID:G172631; PIDN:AA35059.1; PID:G172632
R:Yoshimoto, H.; Yamashita, I.
Mol. Gen. Genet. 228, 270-280, 1991
A:Title: The GML/SNF2 gene of Saccharomyces cerevisiae encodes a highly charged nuclear
A:Reference number: S16820; MUID:91360076; PMID:1886612
A:Accession: S16820
A:Molecule type: DNA
A:Residues: 1-1703 <YOS>
A:Cross-references: EMBL:X57837; NID:G4499; PIDN:CAA0969.1; PID:G4500
R:Cheret, G.; Sor, F.
Submitted to the Protein Sequence Database, July 1996
A:Reference number: S67169
A:Accession: S67192
A:Molecule type: DNA
A:Residues: 1-1703 <CHE>

A:Cross-references: EMBL:Z75198; NID:g1420643; PIDN:CAA9517.1; PID:e252424; PID:g1420644
A:Experimental source: strain S288C
R:Gizlupich, C.; Janniaux, J.C.; Kordae, E.; Poirey, R.; Pujol, A.; Tobiasch, E.
Submitted to the Protein Sequence Database, July 1996
A:Reference number: S67194
A:Accession: S67194

A:Molecule type: DNA
A:Residues: 1-108 <CZI>
A:Cross-references: EMBL:Z75198; MIPS:YOR290C
A:Experimental source: strain S288C
R:Cheret, G.; Bernardi, A.; Sor, F.

Yeast 12, 1059-1064, 1996
A:Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces
A:Reference number: S72039; MUID:97051594; PMID:8896271
A:Accession: S72058

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1703 <CHW>
A:Cross-references: EMBL:X89633; NID:g1279694; PIDN:CAA61793.1; PID:e189413; PID:g127971

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Gene: SGD:SNF2; GML1
A:Cross-references: MIPS:YOR290C; SGD:S0005816
A:Map position: 15R

C:Superfamily: unassigned bromodomain proteins; bromodomain homology

C:Keywords: nucleus; transcription regulation

F:1576-1631/Domain: bromodomain homology <BRO>

Qy 1 SHQESTGRSRGRSGXSG 18
Db 1499 SARTSTRGRGRGRGRG 1516

RESULT 15

T15092
Hypothetical protein K06A5.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15092
R:Wamsley, P.
Submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid K06A5.
A:Reference number: Z18291
A:Accession: T15092
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-662 <WAM>
A:Cross-references: EMBL:AF039018; NID:g2736359; PID:g2736362; PIDN:AA894170.1; GSPDB:GNT
A:Experimental source: strain Bristol N2; clone K06A5
C:Genetics:
A:Gene: CESP:K06A5.1
A:Map position: 1
A:Introns: 5/3; 47/3; 157/3; 509/3; 590/3; 646/3

Query Match 48.4%; Score 44.5; DB 2; Length 682;
Best Local Similarity 61.1%; Pred. No. 35;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy 5 STRGRSRGRSGXSGS 19
Db 358 TTRSRGRGRFELSGNSGA 375

Search completed: September 28, 2004, 06:15:00
Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQESTRGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	91.3	416	1	FILIA_HUMAN
2	58	63.0	506	1	VE2_HPV47
3	48	52.2	772	1	MR11_CAEEL
4	47	51.1	308	1	GRP3_ARTSA
5	45	48.9	70	1	L2W0_ADE40
6	45	48.9	1703	1	SNF2_YEAST
7	44.5	48.4	697	1	AN3_XENLA
8	44	47.8	452	1	VE2_HPV17
9	44	47.8	483	1	VE2_HPV14
10	44	47.8	734	1	ZN45_HUMAN
11	44	47.8	1232	1	Y005_CAEEL
12	43	46.7	123	1	LSM4_CAEEL
13	43	46.7	210	1	Y043_HA1N1
14	43	46.7	239	1	U2AG_MOUSE
15	43	46.7	240	1	U2AG_MOUSE
16	43	46.7	488	1	VE2_HPV49
17	43	46.7	514	1	VE2_HPV05
18	43	46.7	774	1	MRK2_MOUSE
19	43	46.7	2233	1	COAC_YEAST
20	42	45.7	184	1	TRSF_DROSI
21	42	45.7	197	1	TRSF_DROME
22	42	45.7	694	1	FRZ2_DROME
23	42	45.7	706	1	MR11_MOUSE
24	42	45.7	1015	1	PRQ_CRASP
25	42	45.7	2404	1	SON_MOUSE
26	42	45.7	2426	1	SON_HUMAN
27	41.5	45.1	1180	1	TYK2_MOUSE
28	41	44.6	176	1	SSB_TREPA
29	41	44.6	197	1	RX21_DROME
30	41	44.6	315	1	SOL2_YEAST
31	41	44.6	325	1	S265_MOUSE
32	41	44.6	332	1	S265_RAT
33	41	44.6	337	1	Z265_HUMAN

34	41	44.6	503	1	VE2_HPV21	P50767 human papil
35	41	44.6	525	1	MR11_RAT	P48965 rattus norv
36	41	44.6	526	1	ZN45_YEAST	Q00362 saccharomyc
37	41	44.6	655	1	ILF1_HUMAN	Q01167 homo sapien
38	41	44.6	706	1	MR11_RAT	Q91160 rattus norv
39	41	44.6	708	1	MR11_HUMAN	P49955 homo sapien
40	41	44.6	825	1	S55_RAT	Q63003 rattus norv
41	41	44.6	1061	1	PRD4_HUMAN	Q94966 homo sapien
42	41	44.6	1337	1	DEXT_STRDO	P39653 streptococc
43	40	43.5	148	1	LSM4_FAGSY	Q92809 fagus sylv
44	40	43.5	206	1	BTJ3_HUMAN	P20290 homo sapien
45	40	43.5	255	1	THO4_MOUSE	O08583 mus musculu

ALIGNMENTS

```

RESULT 1
FILIA_HUMAN          STANDARD;          PRT;          416 AA.
ID      P20930;
AC      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Filaggrin precursor (Fragment).
GN      FLG.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89296901; PubMed=2740331;
RA      McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA      Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RT      "Characterization of a cDNA clone encoding human filaggrin and
RT      localization of the gene to chromosome region 1q21.";
RL      Proc. Natl. Acad. Sci. U.S.A. 86:4846-4852(1989).
RN      [2]
RP      CITRULLINATION.
RX      MEDLINE=9637438; PubMed=8780679;
RA      Senoh T., Kan S., Ogawa H., Manabe M., Asaga H.;
RT      "Preferential desimination of keratin K1 and filaggrin during the
RT      terminal differentiation of human epidermis.";
RL      Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC      -!- FUNCTION: Aggregates Keratin intermediate filaments and promotes
CC      disulfide-bond formation among the intermediate filaments during
CC      terminal differentiation of mammalian epidermis.
CC      -!- PTM: Filaggrin is initially synthesized as a large, insoluble,
CC      highly phosphorylated precursor containing many tandem copies of
CC      324 AA, which are not separated by "large linker". The precursor
CC      is deposited as keratohyalin granules. During terminal
CC      differentiation it is dephosphorylated and proteolytically
CC      cleaved.
CC      -!- PTM: Undergoes deimination of some arginine residues
CC      (citrullination).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; M24355; AA52454.1; -.
DR      PIR; A32947; A32947.
DR      Genew; HGNC:3748; FLG.
DR      MIM; 135940; -.
DR      GO; GO:0005882; C:intermediate filament; NMS.
DR      GO; GO:0005198; F:structural molecule activity; NMS.
DR      GO; GO:0007275; P:development; NMS.
DR      InterPro; IPR003503; Filaggrin.

```

DR Pfam; PF03516; Flaggrin; 2.
DR PRINTS; PR00487; FLAGGRIN.
KW Phosphorylation; Citrullination; Developmental protein.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44105 MW; DEEA218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
Best Local Similarity 89.5%; Pred. No. 6.3e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SH0ESTGRSRGRSGXSGS 19
Db 7 SH0ESTGRSRGRSGXSGS 25

RESULT 2

VE2_HPV47 STANDARD; PRT; 506 AA.
AC P22420.
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human Papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2RE/S EITHER
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STRICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -.
DR PIR; D5324; W2WL47.
DR HSSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C; 1.
DR Pfam; PF00508; E2_N; 1.
DR Pfam; PF000672; E2_C; 1.
DR Pfam; PF000678; E2_N; 1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF875065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.089;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SH0ESTGRSRGRSGXSGS 19

Db 342 SREGNTRGRGRGRCGAGS 360

RESULT 3

MRE11_CAEEL STANDARD; PRT; 772 AA.
ID MRE11_CAEEL
AC Q23255;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Double-strand break repair protein mre-11.
GN MRE-11 OR ZC302.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP GUD-139.
RX MEDLINE=21135651; PubMed=11238374;
RA Chin G.M., Villeneuve A.M.;
RT "C. elegans mre-11 is required for meiotic recombination and DNA
RT repair but is dispensable for the meiotic G(2) DNA damage
RT checkpoint.";
RL Genes Dev. 15:522-534(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kelly P.F.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
CC Possesses single-strand endonuclease activity and double-strand-
CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
CC processing.
CC -1- FUNCTION: Required for meiotic crossing over and chiasma
CC formation. Pachytene morphology and homolog pairing are normal.
CC Vital in long term for maintenance of reproductive capacity of
CC subsequent generations.
CC -1- COPACITOR: Manganese (By similarity).
CC -1- SUBUNIT: Forms a complex with rad-50 (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the MRE11/RAD32 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73978; CAA98292.1; -.
DR PIR; T27512; T27512.
DR GeronOnline; 208612; -.
DR WormRep; ZC302.1; CE06573.
DR InterPro; IPR003701; DNA_repair.
DR InterPro; IPR004843; M-psestrase.
DR InterPro; IPR007281; Mre11_DNA_bind.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF04152; Mre11_DNA_bind; 1.
DR TIGRfam; TIGR00583; mre11; 1.
KW DNA repair; Hydrolyase; Nuclease; Endonuclease; Exonuclease;
KW Nuclear protein; Manganese; Meiosis.
FT MUTAGEN 139 E->K; IN MRE-11-ME41, DEFECTIVE IN
FT MEIOTIC CHROMOSOME DEGRADATION.
SQ SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;

Query Match 52.2%; Score 48; DB 1; Length 772;
Best Local Similarity 52.6%; Pred. No. 5.1;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 SH0ESTGRSRGRSGXSGS 19

Db 697 SKOPTTRGRGRGARGAGAS 715

RESULT 4

GRP3_ARTSA STANDARD; PRT; 308 AA.
 ID GRP3_ARTSA STANDARD; PRT; 308 AA.
 AC P13230;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycine-rich protein GRP3.
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 NCBI_TaxID=85549;
 RX MEDLINE=88007550; PubMed=2443491;
 RA Cruz-Alvarez M., Pellicer A.;
 RT "Cloning of a full-length complementary DNA for an Artemia salina
 RT glycine-rich protein. Structural relationship with RNA binding
 RT proteins.";
 RL J. Biol. Chem. 262:13377-13380(1987).
 CC -1- PTM: THE ARGININES IN THE GLY-RICH DOMAIN MIGHT BE METHYLATED.
 CC -1- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 CC HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 CC STRUCTURALLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC -1- SIMILARITY: Contains 1 KH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03453; AAC83400.1; -.
 DR PIR: A29379; A29379.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR SMART: SM00322; KH; 1.
 DR PROSITE: PS50084; KH TYPE 1; 1.
 DR Nuclear protein; Ribonucleoprotein; Methylation; RNA-binding.
 FT DOMAIN 83 118 KH.
 FT 186 308 GLY-RICH.
 SQ SEQUENCE 308 AA; 35034 MW; 5534CEE552BB3BC2 CRC64;
 Query Match SNF2 YEAST 51.1%; Score 47; DB 1; Length 308;
 Best Local Similarity 75.0%; Pred. No. 2.7;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 5
 L2MU_ADE40 STANDARD; PRT; 70 AA.
 ID L2MU_ADE40 STANDARD; PRT; 70 AA.
 AC Q64858; Q67171;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Late L2 mu core protein precursor (pmu) (Protein X).
 GN px.
 OS Human adenovirus type 40.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 NCBI_TaxID=28284;
 RX MEDLINE=91187657; PubMed=1901413;
 RA Laurent B.C., Treitel M.A., Carlson M.;
 RT "Functional interdependence of the yeast SNF2, SNF5, and SNF6
 RT proteins in transcriptional activation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2687-2691(1991).
 RC SEQUENCE FROM N.A.
 RC STRAIN=Dugan;
 RX MEDLINE=94087748; PubMed=8263936;
 RA Davison A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
 RT "The DNA sequence of adenovirus type 40.";
 RL J. Mol. Biol. 234:1308-1316(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dugan;
 RX MEDLINE=96332521; PubMed=8760431;
 RA Grydskul J.D., Fortes E., Petric M., Brown M.;
 RT "Common epitope on protein VI of enteric adenoviruses from subgenera
 RT A and F.";
 RL J. Gen. Virol. 77:1811-1819(1996).
 CC -1- FUNCTION: The role of the precursor might be to condense the viral
 CC prochromatin for encapsidation by virtue of the two basic domains
 CC (by similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L19443; AAC13965.1; -.
 DR EMBL: U14651; AAB19000.1; -.
 DR InterPro: IPR008393; Adenovirus_PX.
 DR Pfam: PF05829; Adenovirus_PX; 1.
 DR Core protein; DNA-binding; Late protein.
 FT PROCEP 1 26 BY SIMILARITY.
 FT CHAIN 27 40 LATE L2 MU CORE PROTEIN.
 FT PROCEP 41 70 BY SIMILARITY.
 FT SITE 26 27 CLEAVAGE (BY ADENOVIRUS PROTEASE)
 FT SITE 40 41 CLEAVAGE (BY ADENOVIRUS PROTEASE)
 FT SITE (POTENTIAL).
 FT CONFLICT 31 32 RA -> P (IN REF. 2).
 FT SEQUENCE 70 AA; 7596 MW; BF621E01518PF69 CRC64;
 Query Match SNF2 YEAST 48.9%; Score 45; DB 1; Length 70;
 Best Local Similarity 66.7%; Pred. No. 1.1;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RC STRAIN=AH22; PubMed=1886612;
 RX MEDLINE=91360076; PubMed=1886612;
 RA Yoshimoto H., Yamashita I.;
 RT "The GAM1/SNF2 gene of *Saccharomyces cerevisiae* encodes a highly
 RT charged nuclear protein required for transcription of the STAI
 RT gene";
 RL Mol. Gen. Genet. 228:270-280(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1B;
 RX MEDLINE=95332261; PubMed=7608126;
 RA Kodaki T., Hosaka K., Nakawa J., Yamashita S.;
 RT "The SNF2/SWI2/GAM1/TRY3/RIC1 gene is involved in the coordinate
 RT regulation of phospholipid synthesis in *Saccharomyces cerevisiae*";
 RL J. Biochem. 117:362-368(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97051594; PubMed=8896271;
 RA Cheret G., Bernardi A., Sor F.J.;
 RT "DNA sequence analysis of the VP11-SNF2 region on chromosome XV of
 RT *Saccharomyces cerevisiae*";
 RL Yeast 12:1059-1064(1996).
 RN [5]
 RP SEQUENCE OF 1-309 FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97298310; PubMed=9153758;
 RA Poljey R., Ciepluch C., Tobiasch E., Pujol A., Kordes E.,
 RA Janniaux J.-C.;
 RT "Sequence and analysis of a 36.2 kb fragment from the right arm of
 RT yeast chromosome XV reveals 19 open reading frames including SNF2 (5'
 RT end), CPA1, SLY4, a putative transport ATPase, a putative ribosomal
 RT protein and an SNF2 homologue";
 RL Yeast 13:479-482(1997).
 RN [6]
 RP MUTAGENESIS.
 RX MEDLINE=97025355; PubMed=8871545;
 RA Richmond E., Peterson C.L.;
 RT "Functional analysis of the DNA-stimulated Arpase domain of yeast
 RT Swi2/SNF2";
 RL Nucleic Acids Res. 24:3685-3692(1996).
 CC -1- FUNCTION: Involved in transcriptional activation. The SWI/SNF
 CC complex is required for the induced expression of a large number
 CC of genes. This complex alters chromatin structure to facilitate
 CC binding of gene-specific dedicated transcription factors.
 CC -1- SUBUNIT: Component of the SWI/SNF global transcription activator
 CC complex.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: STRONG, TO DROSOPHILA BRAHMA.
 CC -1- SIMILARITY: Contains 3 A.T hook DNA-binding repeats.
 CC -1- SIMILARITY: Contains 3 A.T hook DNA-binding repeats.
 CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M61703; AAA5059.1; -;
 DR EMBL: X57837; CAA40969.1; -;
 DR EMBL: D90459; BAA14423.1; -;
 DR EMBL: X89633; CAA61793.1; -;
 DR EMBL: Z75198; CAA9517.1; -;
 DR EMBL: Z75199; CAA9519.1; -;
 DR PIR: S15047; S15047; -;
 DR GenBank: 143877; -;
 DR TRANSFAC: T02401; -;
 DR SCD: S0005815; SNF2; -;
 DR InterPro: IPR000637; AT_hook.
 DR InterPro: IPR001487; Bromodomain.

DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001550; Helicase_C.
 DR InterPro: IPR000330; SNF2_N.
 DR Pfam: PF02178; AT_hook_2.
 DR Pfam: PF00439; bromodomain_1.
 DR Pfam: PF00271; helicase_C_1.
 DR Pfam: PF00176; SNF2_N_1.
 DR PRINTS: PR00929; ATHOOK.
 DR PRINTS: PR00503; BROMODOMAIN.
 DR SMART: SM00384; AT_hook_2.
 DR SMART: SM00297; BROMO_1.
 DR SMART: SM00487; DEHDC_1.
 DR SMART: SM00490; HELIC_C_1.
 DR PROSITE: PS00633; BROMODOMAIN_1; 1.
 DR PROSITE: PS50014; BROMODOMAIN_2; 1.
 KW Transcription regulation; Nuclear protein; Activator; Repeat;
 KW Bromodomain; ATP-binding; Helicase.
 FT DOMAIN 55
 FT NP BIND 207 239 ALA/GLN-RICH.
 FT SITE 792 799 ATP (BY SIMILARITY).
 FT DNA BIND 1446 1456 DEH BOX.
 FT DNA BIND 1502 1513 A.T HOOK 1.
 FT DNA BIND 1516 1526 A.T HOOK 2.
 FT DOMAIN 1568 1638 BROMODOMAIN.
 FT SEQUENCE 1703 AA; 194050 MW; 84BB8C55C8F3E6D CRC64;
 SQ
 Query Match 48.9%; Score 45; DB 1; Length 1703;
 Best Local Similarity 55.6%; Pred. No. 35;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHQSTGRGRSGRSGXSG 18
 Db 1499 SARTSTRGRGRCGRGRC 1516
 RESULT 7
 AN_XENLA STANDARD: PRT; 697 AA.
 ID AN3_XENLA
 AC P24346;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative ATP-dependent RNA helicase An3.
 GN AN3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xeropodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91141586; PubMed=1996140;
 RA Gururajan R., Perry O'Keefe H., Melton D.A., Weeks D.L.;
 RT "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
 RT RNA helicase";
 RL Nature 349:717-719(1991).
 CC -1- FUNCTION: Putative ATP-dependent RNA helicase.
 CC -1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X57328; CAA40605.1; -;
 DR PIR: S13654; S13654.
 DR HSSP: Q58083; 1HV8.
 DR InterPro: IPR001410; DEAD.

```

DR InterPro: IPR000629; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD. 1.
DR Pfam: PF00271; Helicase_C. 1.
DR SMART: SM00487; DEXDC. 1.
DR SMART: SM00490; HELICG. 1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE. 1.
KW Helicase; ATP-binding; RNA-binding.
FT NE_BIND 265 272 ATP (BY SIMILARITY).
FT SITE 388 391 DEAD_BOX.
FT DOMAIN 623 697 GLY/SER-RICH.
SQ SEQUENCE 697 AA; 77302 MW; F3DD23EB60B2E2EF CRC64;

Query Match
Best Local Similarity 52.8%; Score 44.5; DB 1; Length 697;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 2 HQSTRGRSRGR-SGXSGS 19
DB 621 HKSSRGRSRGRSGGGA 639

```

```

RESULT 8
ID VE2_HPVL7 STANDARD; PRT; 452 AA.
AC P36785;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 17.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10607;
RN (1)
RP MEDLINE=94265501; Pubmed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -!- SUBUNIT: Binds DNA as a dimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74469; CAAS2515.1; -.
DR PIR: S36482; S36482.
DR HSSP: P17383; IDHM.
DR InterPro: IPR000427; E2_C.
DR InterPro: IPR001866; E2_N.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF00511; E2_C. 1.
DR Pfam: PF00508; E2_N. 1.
DR ProDom: PD000672; E2_C. 1.
DR ProDom: PD000678; E2_N. 1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 452 AA; 51241 MW; 70298AFL6D3ESPS CRC64;

```

```

Query Match
Best Local Similarity 50.0%; Score 44; DB 1; Length 452;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 3 QESTRGRSRGRSGXSG 18
DB 293 RDSRSPNRGRSGSSG 308

```

```

RESULT 9
ID VE2_HPVL4 STANDARD; PRT; 483 AA.
AC P36783;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 14.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10605;
RN (1)
RP MEDLINE=94265501; Pubmed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -!- SUBUNIT: Binds DNA as a dimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74467; CAAS2503.1; -.
DR PIR: S36470; S36470.
DR HSSP: P17383; IDHM.
DR InterPro: IPR000427; E2_C.
DR InterPro: IPR001866; E2_N.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF00511; E2_C. 1.
DR Pfam: PF00508; E2_N. 1.
DR ProDom: PD000672; E2_C. 1.
DR ProDom: PD000678; E2_N. 1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 483 AA; 54938 MW; CFE923ADFAE3800 CRC64;

Query Match
Best Local Similarity 50.0%; Score 44; DB 1; Length 483;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 SHQSTRGRSRGRSGXSG 18
DB 327 SSKRGRGRGRSGGSG 344

RESULT 10
ZNA2_HUMAN

```

ID ZN42 HUMAN STANDARD; PRT; 734 AA.
 AC P28698; Q9NRX0; Q9DWM2;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Zinc finger protein 42 (Myeloid zinc finger 1) (MZF-1).
 GN ZNF42 OR MZF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RX MEDLINE=91317761; PubMed=1860835;
 RP SEQUENCE FROM N.A. (ISOFORM MZF1A).
 RA Hromas R., Collins S.J., Hickstein D., Raskind W., Deaven L.L.,
 RA O'Hara P., Hagen F.S., Kaushansky K.;
 RT "A retinoic acid-responsive human zinc finger gene, MZF-1,
 RT preferentially expressed in myeloid cells.";
 RL J. Biol. Chem. 266:14183-14187 (1991).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM MZF1A AND MZF1B-C).
 RC TISSUE=BONE MARROW; PubMed=10974541;
 RX MEDLINE=20432092; PubMed=10974541;
 RA Peterson M.J., Morris J.F.;
 RT "Human myeloid zinc finger gene MZF produces multiple transcripts and
 RT encodes a SCAN box protein.";
 RL Gene 254:105-118 (2000).
 CC -!- FUNCTION: May be one regulator of transcriptional events during
 CC hemopoietic development.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=MZF1A; Synonyms=MZF1B;
 CC IsoId=P28698-1; Sequence=Displayed;
 CC Name=MZF1B-C;
 CC IsoId=P28698-2; Sequence=VSP_006889, VSP_006890;
 CC -!- TISSUE SPECIFICITY: Preferentially expressed in differentiating
 CC myeloid cells.
 CC -!- INDUCTION: By retinoic acid.
 CC -!- SIMILARITY: BELONGS TO THE KUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: Contains 1 SCAN box domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL; M58297; AAA59898.1; -;
 DR EMBL; AF055078; AAD55810.1; -;
 DR EMBL; AF055077; AAD55809.1; -;
 DR EMBL; AF161886; AAF80466.1; -;
 DR EMBL; AF161886; AAF80465.1; -;
 DR PIR; A40751; A40751.
 DR HSSP; P08047; I5P2.
 DR TRANSFAC; T00529; -;
 DR Genew; HGNC:13108; ZNF42.
 DR MIM; 194550; -;
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR InterPro; IPR003309; Treg SCAN.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF02023; SCAN; 1.
 DR Pfam; PF00096; zf_C2H2; 13.
 DR PRINTS; PR00048; ZINCFINGER.
 DR PRODOM; PD000003; Znf_C2H2; 6.
 DR SMART; SM00431; LER; 1.
 DR SMART; SM00355; ZNF_C2H2; 13.
 DR PROSITE; PS50804; SCAN_BOX; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.

DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 13.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat; Alternative splicing; Polymorphism.
 FT DOMAIN 44 125
 FT 310 321 ASP/GLU-RICH (ACIDIC).
 FT ZN_FING 356 378 C2H2-TYPE.
 FT ZN_FING 384 406 C2H2-TYPE.
 FT ZN_FING 412 434 C2H2-TYPE.
 FT ZN_FING 440 462 C2H2-TYPE.
 FT ZN_FING 463 484 GLY/PRO-RICH.
 FT DOMAIN 485 507
 FT ZN_FING 513 535 C2H2-TYPE.
 FT ZN_FING 541 563 C2H2-TYPE.
 FT ZN_FING 569 591 C2H2-TYPE.
 FT ZN_FING 597 619 C2H2-TYPE.
 FT ZN_FING 625 647 C2H2-TYPE.
 FT ZN_FING 653 675 C2H2-TYPE.
 FT ZN_FING 681 703 C2H2-TYPE.
 FT ZN_FING 709 731 C2H2-TYPE.
 FT VARSPLIC 1 249 Missing (in isoform MZF1B-C).
 FT VARSPLIC 250 257 /FTId=VSP_006889.
 FT VARSPLIC 257 /FTId=VSP_006890.
 FT VARIANT 331 331 I -> V (in dbSNP:4756).
 FT CONFLICT 304 305 /FTId=VAR_014826.
 FT SEQUENCE 734 AA; 82036 MW; 2BE7D69B18F29437 CRC64;
 SQ
 Query Match 47.8%; Score 44; DB 1; Length 734;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 5 STRGRSGRSGXSG 18
 Db 337 STRGRSGRSPGSG 350
 RESULT 11
 ID Y005 CAEEL STANDARD; PRT; 1232 AA.
 AC P34613;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK512.5 in chromosome III.
 GN ZK512.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berts M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkes T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kerhaw J., Kirsten J., Lalister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J.,
 RA Weidman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38 (1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: Z23177; CA80146.1; -.
DR PIR: S40766; S40766.
DR WormPep; ZK512.5; CE00411.
KW Hypothetical protein.
SQ SEQUENCE 1232 AA; 134923 MW; 6DFC35D664AA8D6A CRC64;

Query Match 47.8%; Score 44; DB 1; Length 1232;
Best Local Similarity 64.3%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HOESTRGRSRGRSG 15
Db 110 HONSSRGPSGFGSG 123

RESULT 12
LSM4_CAEEL STANDARD; PRT; 123 AA.
AC Q19953;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.
GN F32A5.7.
OS Caenorhabditis elegans.
OC Rhabditiidae; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditiodea;
OC Rhabditiidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (by similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (by
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: U20864; AAC4661.1; -.
DR PIR: T16234; T16234.
DR WormPep; F32A5.7; CE01277.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00551; Sm; 1.
DR SMART; SM00551; Sm; 1.
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
KW RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTGRSRGRSGXSG 18
Db 93 REOSRGRGGRGGRG 108

```

```

RESULT 13
Y043 HALN1 STANDARD; PRT; 210 AA.
AC 09NH82;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Yng2543c.
GN YNG2543C.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Sprogon J.,
RA Swartzell S., Weir D., Hall U., Dahl T.A., Weir D., Goo Y.A.,
RA Letchauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs W.P., Angvine C.M., Dale H.,
RA Isebnarger T.A., Peck R.F., Pohlechröder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RA "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- SIMILARITY: Contains 1 AMMECR1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: AB005130; AAG20599.1; -.
DR PIR: C64404; C64404.
DR HAMAP; MF_00645; atypical; 1.
DR InterPro; IPR002733; DUF51.
DR Pfam; PF01871; AMMECR1; 1.
DR ProDom; PD009671; DUF51; 1.
DR TIGRFAMs; TIGR00296; TIGR00296; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 210 AA; 22964 MW; D4CBEDCEB4307EFF CRC64;

Query Match 46.7%; Score 43; DB 1; Length 210;
Best Local Similarity 61.1%; Pred. No. 76;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 4 ESTRGRS--RGRSGXSGS 19
Db 55 ESTRGRRLRGCGAAGS 72

RESULT 14
U2AG_MOUSE STANDARD; PRT; 239 AA.
ID U2AG_MOUSE
AC U2AG83; O99LX2; O9C298;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE subunit) (U2 snRNP auxiliary factor small subunit).
GN U2AF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;

```


RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Teshiyuki S., Carninci P., Parise C.,
RA Raha S.S., Locantello N.A., Peters G.J., Abramson R.D., Molligh S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy U., Helton E., Keltman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Stailus D.E.,
RA Schnerch A., Schein J.E., Jones S.U.M., Maria W.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP FUNCTION
RX MEDLINE=96249383; PubMed=8647433;
RA Zhu P., Maniatis T.;
RT "The splicing factor U2AF35 mediates critical protein-protein
RT interactions in constitutive and enhancer-dependent splicing.";
RN Genes Dev. 10:1356-1368(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 43-146 IN COMPLEX WITH U2AF2,
RP AND MUTAGENESIS OF TRP-134.
RX MEDLINE=21435806; PubMed=11551507;
RA Klempf C.L., Rodionova N.A., Green M.R., Buxley S.K.;
RT "A novel peptide recognition mode revealed by the X-ray structure of
RT a core U2AF35/U2AF65 heterodimer.";
RL Cell 106:595-605(2001).
CC -!- FUNCTION: Plays a critical role in both constitutive and enhancer-
CC dependent splicing by mediating protein-protein interactions and
CC protein-RNA interactions required for accurate 3' splice site
CC selection. Recruits U2 snRNP to the branch point. Directly
CC mediates interactions between U2AF2 and proteins bound to the
CC enhancers and thus may function as a bridge between U2AF2 and the
CC enhancer complex to recruit it to the adjacent intron.
CC -!- SUBUNIT: Heterodimer with U2AF2.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: The C-terminal SR-rich domain is required for
CC interactions with SR proteins and the splicing regulators TRA and
CC TRA2, and the N-terminal domain is required for formation of
CC the U2AF1/U2AF2 heterodimer.
CC -!- SIMILARITY: Belongs to the SR family of splicing factors.
CC -!- SIMILARITY: Contains 2 C3H1-type zinc fingers.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M69682; AAA36619.1; -;
DR EMBL: AP001748; BAA95534.1; -;
DR EMBL: BC001177; AA01177.1; -;
DR EMBL: BC001923; AA01923.1; -;
DR PIR: A46179; A46179.
DR PDB: 1UMT; 19-SEP-01.
DR Genew: HGNC:12453; U2AF1.
DR GK: 001081; -;
DR KIM: 191317; -;
DR GO: GO:0015030; C:capal body; TAS.
DR GO: GO:0008248; F:pre-mRNA splicing factor activity; TAS.
DR GO: GO:0006397; F:mRNA processing; TAS.
DR GO: GO:0006371; P:mRNA splicing; TAS.
DR InterPro: IPR005054; RNA_rec_mot.
DR InterPro: IPR005571; Znf_CCH.

DR Pfam: PF00076; rtm; 1.
DR Pfam: PF00642; zf-CCCH; 2.
DR SMART: SM00360; RRM; 1.
DR SMART: SM00356; Znf_C3H1; 2.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE NEG.
KW mRNA processing; mRNA splicing; Nuclear protein; Spliceosome;
KW RNA-binding; Metal-binding; Zinc; Repeat; Zinc-finger; 3D-structure.
FT ZN_FING 13 41
FT C3H1-TYPE 1.
FT DOMAIN 65 147
FT ZN_FING 149 173
FT DOMAIN 178 240
FT C3H1-TYPE 2.
FT ARG/GLY/SER-RICH (RS DOMAIN).
FT POLY-GLY.
FT MUTAGEN 134 134 W->A: DECREASES AFFINITY FOR UAF2 BY 3
FT ORDERS OF MAGNITUDE.
SQ SEQUENCE 240 AA; 27872 MW; 3DA130DC0B953F6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 240;
Best Local Similarity 44.4%; Pred. No. 8.7;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTGRSRGRGSG 18
DB 199 SRERSRGRDRGRGGGG 216

Search completed: September 28, 2004, 06:05:49
Job time : 8.22396 secs

Blank sheet


```

015206
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814]";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10FA91991 CRC64;

Query Match 93.5%; Score 86; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 2.4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
DB 551 SHOESTRGRSGRSGSGS 569

RESULT 3
005331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipkev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 34 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1; -.
DR PIR; A4818; A4818.
DR HSSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CAP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 93.5%; Score 86; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
DB 773 SHOESTRGRSGRSGSGS 791

RESULT 4
Q9H4U3 PRELIMINARY; PRT; 798 AA.
AC Q9H4U3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJIAN1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; Filaggrin.
DR NON_TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDAD8D1290805 CRC64;

Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
DB 428 SHOESTRGRSGRSGSGS 446

RESULT 5
Q01212 PRELIMINARY; PRT; 1084 AA.
ID Q01212

```

AC 001212; C03840;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9106347; PubMed=2248957;
 RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]."
 RT 11;30(23):5814.";
 RL Biochemistry 29:9432-9440(1990).
 DR EMBL; M60503; AAA63243.1; JOINED.
 DR EMBL; M60501; AAA63243.1; JOINED.
 DR GO; GO:0005882; C:intermediate filament; NAS.
 DR GO; GO:0005198; F:structural molecule activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PRO0487; Filaggrin.
 FT NON TER
 FT SEQUENCE 1084 AA; 115271 MW; 80046408BD5A362D CRC64;

Query Match 91.3%; Score 84; DB 4; Length 1084;
 Best Local Similarity 89.5%; Pred. No. 5.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 64 SHOESTRGRSGRSGSGS 82

RESULT 6
 ID 001720 PRELIMINARY; PRT; 591 AA.
 AC 001720;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Filaggrin precursor (PROFILAGGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Presland R.B., Haycock P.V., Fleckman P., Nirmuskesiri W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 organization and identification of an S-100-like calcium binding
 domain at the amino terminus."
 RT J. Biol. Chem. 267:23772-23781(1992).
 CC -I- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS DURING
 DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -I- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 324 AA. THE PRECURSOR IS DEPOSITED AS KERSOTHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTOLYTICALLY CLEAVED.
 CC -I- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; L01089; AAA60177.1; -;
 DR EMBL; L01090; AAA60176.1; -;

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; Filaggrin.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KW Polymorphism.
 FT PROPEP 1 293
 FT CHAIN 294 467
 FT PROPEP 468 474
 FT CHAIN 475 >591
 FT CA BIND 19 32
 FT CA BIND 62 73
 FT NON TER 591 591
 FT SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 591;
 Best Local Similarity 84.2%; Pred. No. 0.00084;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 449 SHOESTRGRSGRSGSGS 467

RESULT 7
 ID 09H4U2 PRELIMINARY; PRT; 687 AA.
 AC 09H4U2;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Df14N1.1.1 (Profilaggrin 5' end) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL356504; CAC13172.1; -;
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; Filaggrin.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON TER 687 687
 FT SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 687;
 Best Local Similarity 84.2%; Pred. No. 0.00098;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 449 SHOESTRGRSGRSGSGS 467

RESULT 8

ID Q03838 PRELIMINARY; PRT; 465 AA.

AC Q03838;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGRIN (PROFILAGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

FT NON_TER

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

REVISIONS.

RA MEDLINE=51255199; PubMed=2043621;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RT Biochemistry 29:9432-9440(1990).

RL [2]

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND

CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND

CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

CC REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,

CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin

CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.

CC EMBL; M60499; AAA63246.1; -;

CC GO; GO:0005198; F:structural molecule activity; IEA.

CC InterPro: IPR003303; Filaggrin.

CC Pfam; PF03516; Filaggrin; 3.

CC PRINTS; PR00487; FILAGRIN.

CC NON_TER

CC SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

Query Match

Best Local Similarity 76.1%; Score 70; DB 4; Length 465;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19

DB 227 SHOESTRGRSGRSGSGS 245

RESULT 9

ID Q75370 PRELIMINARY; PRT; 322 AA.

AC Q75370;

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Epidermal filaggrin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

FT NON_TER

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

REVISIONS.

RA MEDLINE=51255199; PubMed=2043621;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RT Biochemistry 29:9432-9440(1990).

RL [2]

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND

CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND

CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

CC REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,

CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin

CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.

CC EMBL; M60499; AAA63246.1; -;

CC GO; GO:0005198; F:structural molecule activity; IEA.

CC InterPro: IPR003303; Filaggrin.

CC Pfam; PF03516; Filaggrin; 3.

CC PRINTS; PR00487; FILAGRIN.

CC NON_TER

CC SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

OX NCBI_TaxID=9606;

FT NON_TER

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

REVISIONS.

RA MEDLINE=59101527; PubMed=9886436;

RA Girbal-Neuhausser E., Durieux J.J., Arnaud M., Dalbon P., Sebbaq M.,

RA Vincent C., Simon M., Senshu T., Masson-Besiere C.,

RA Jolivet-Reynaud C., Delivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated

RT anti-filaggrin autoantibodies are posttranslationally generated on

RT various sites of (pro)filaggrin by deamination of arginine residues.";

RT J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC23559.1; -;

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER

FT NON_TER

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

Query Match

Best Local Similarity 75.0%; Score 69; DB 4; Length 322;

Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 18

DB 305 SHOESTRGRSGRSGSGS 322

RESULT 10

ID Q80890 PRELIMINARY; PRT; 476 AA.

AC Q80890;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE EBNA1.

OS Herpesvirus papio.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gammaherpesvirinae; Lymphocryptovirus.

CC NCBI_TaxID=10394;

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN=594-S;

CC MEDLINE=96400106; PubMed=8806482;

CC Yates J.L., Camilo S.M., Ali S., Ying A.;

CC "Comparison of the EBNA1 proteins of Epstein-Barr virus and

CC herpesvirus papio in sequence and function.";

CC Virology 222:1-13(1996).

CC EMBL; U23857; AAA6373.1; -;

CC HSP; P03211; IVH1.

CC GO; GO:0042025; C:host cell nucleus; IEA.

CC GO; GO:0003677; F:DNA binding; IEA.

CC InterPro: IPR004186; EBNA1.

CC InterPro: IPR009021; Viral_DNA_Db.

CC Pfam; PF02905; EBNA1; 1.

CC SEQUENCE 476 AA; 46082 MW; 3F598F27919F7DA9 CRC64;

Query Match

Best Local Similarity 56.5%; Score 52; DB 12; Length 476;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRGRSGRSGXSGS 19

DB 160 SRGRSGRSGXSGS 173

RESULT 11

ID Q8H8P6 PRELIMINARY; PRT; 188 AA.

AC Q8H8P6;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)


```

DE 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Hypothetical protein.
GN OSUNBA0057E11.19.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarioidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Guyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsaltrin T., Kim M.M., Bera J.U., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Heiso J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uterback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.C., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0057E11 genomic sequence.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC091234; AAN64459.1; -
DR InterPro: IPR008552; DUF834.
DR Pfam: PF05754; DUF834; 1.
KM Hypothetical protein.
SQ SEQUENCE 188 AA; 19599 MW; DCSA0096F51192B5 CRC64;

Query Match 52.2%; Score 48; DB 10; Length 188;
Best Local Similarity 60.0%; Pred. No. 5;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 ESTGSRGSRGSGXG 18
Db 168 EQIRGERGKSGXG 182

RESULT 12
Q943T8 PRELIMINARY; PRT; 336 AA.
ID Q943T8;
AC Q943T8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE B1158F07.1 protein (OJ1008_F01.12 protein).
GN B1158F07.1 OR OJ1008_F01.12.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarioidae; Oryzae; Oryza.
OX NCBI_TaxID=4530; 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=O. sativa; STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
RT clone:OJ1008_F01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=O. sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
RT clone:OJ1008_F01.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003208; BAB64590.1; -
DR EMBL: AP004320; BAC00705.1; -
DR Gramene: Q943T8; -
DR InterPro: IPR008552; DUF834.
DR Pfam: PF05754; DUF834; 1.
SQ SEQUENCE 336 AA; 35999 MW; 5080CDAFACTB9B88 CRC64;

Query Match 52.2%; Score 48; DB 10; Length 336;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 ESTGSRGSRGSGXG 18
Db 316 EQIRGERGKSGXG 330

RESULT 13
Q69375 PRELIMINARY; PRT; 820 AA.
ID Q69375;
AC Q69375;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Putative UL25 protein.
OS Mouse cytomegalovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10366;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K181;
RX MEDLINE=94233727; PubMed=7513920;
RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;
RT "Identification and characterization of a murine cytomegalovirus gene
RT with homology to the UL25 open reading frame of human
RT cytomegalovirus.";
RL Virology 200;643-650(1994).
DR EMBL: U02500; AAL19449.1; -
DR InterPro: IPR006731; Herpes_PP85.
DR Pfam: PF04637; Herpes_PP85; 1.
SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;

Query Match 52.2%; Score 48; DB 12; Length 820;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QESTGSRGSRGSRG 14
Db 186 QSTAGSRGSRGSRG 197

RESULT 14
Q9P312 PRELIMINARY; PRT; 822 AA.
ID Q9P312;
AC Q9P312;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12P1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL390091; CAB98213.1; -
DR PIR: T51049; T51049.
DR GO: GO:0003676; P:nucleic acid binding; IEA.
DR InterPro: IPR005054; RNA_rec_mot.
DR InterPro: IPR005120; Smg-4_UF3.
DR Pfam: PF00076; rrm; 1.
DR Pfam: PF03467; Smg4_UF3; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.

```

SO SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;
 Query Match 52.2%; Score 48; DB 3; Length 822;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTGRSGRSGSGS 19
 :|||:|||||:
 Db 414 RESASGRTRGRGRGT 430

RESULT 15

Q86ZHL PRELIMINARY; PRT; 1229 AA.
 ID Q86ZHL
 AC Q86ZHL; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Related to cell division cycle 2-related protein kinase 7.
 GN 7F4.060.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Newes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BX294020; CAD70910.1; -;
 DR GO: GO:0005524; P:ATP binding; IEA.
 DR GO: GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO: GO:0000910; P:cytokinesis; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR InterPro: IPR008271; Ser Thr kinase.
 DR InterPro: IPR01245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD00001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYRKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Cell division; Kinase.
 SQ SEQUENCE 1229 AA; 136776 MW; 43BE612656FF73E4 CRC64;
 Query Match 52.2%; Score 48; DB 3; Length 1229;
 Best Local Similarity 57.9%; Pred. No. 42;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 :|||:|||||:
 Db 242 SHHRRRRSDKRRSGRSRS 260

Search completed: September 28, 2004, 06:12:44
 Job time : 42.1771 secs

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 1 SHOESTRGRSRGRSGSGS 19

RESULT 2

AAW61514
 ID AAW61514 standard; peptide; 19 AA.

XX AAW61514;

XX 26-OCT-1998 (first entry)

XX Peptide cf, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;

Best Local Similarity 94.7%; Pred. No. 1.7e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 1 SHOESTRGRSRGRSGSGS 19

RESULT 3

AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

XX 26-OCT-1998 (first entry)

XX Peptide cfQ, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 93.5%; Score 86; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 7.5e-07;

Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 1 SHOESTRGRSRGRSGSGS 19

DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 4

ABB97605
 ID ABB97605 standard; protein; 1467 AA.

XX ABB97605;

XX 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 873.

XX Human; antianemic; vulnery; antiinflammatory; immunomodulator;

XX antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;

XX neuroprotective; antiparkinsonian; protein therapy; BSR;

XX expressed sequence tag.

XX Homo sapiens.

OS

```

XX PN W0200222660-A2.
XX PD 21-MAR-2002.
XX PF 10-SEP-2001; 2001WO-US026015.
XX PR 11-SEP-2000; 2000US-00659671.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Dermanac RT;
XX DR WPI; 2002-282408/33.
XX DR N-PSDB; ABN32791.
XX PT An isolated polynucleotide for treating diseases associated with its
XX PT encoded polypeptide such as cancer and multiple sclerosis.
XX PS Example 2; SEQ ID NO 873; 509pp; English.
XX CC The present invention provides the protein and coding sequences of 444
XX CC novel human proteins. These were isolated from expressed sequences tags
XX CC (ESTs). They can be used to stimulate cell growth, to regulate
XX CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX CC rheumatoid arthritis, and to treat nervous system disorders e.g.
XX CC Parkinson's disease. The present sequence is a protein of the invention
XX SO Sequence 1467 AA;

Query Match 93.5%; Score 86; DB 5; Length 1467;
Best Local Similarity 89.5%; Pred. No. 6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
   ||||| ||||| ||||| |||||
Db 773 SHOESTRGRSRGRSGSGS 791

RESULT 5
AAW61516
ID AAW61516 standard; peptide; 19 AA.
XX AAW61516;
XX AC
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cfe, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX OS Synthetic.
XX OS Homo sapiens.
XX PN W09822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

```

```

XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX SQ Sequence 19 AA;

Query Match 92.4%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
   ||||| ||||| ||||| |||||
Db 1 SHOESTRGRSRGRSGSGS 19

RESULT 6
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX AAW61506;
XX AC
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 9 /note= "Citrulline"
XX PN W09822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 7
 ID AAW61508
 AC AAW61508; standard; peptide; 19 AA.

XX 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "Citrulline"

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 FT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 8

ID AAW61515
 AC AAW61515; standard; peptide; 19 AA.

XX 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 FT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 9

ID AAW61507
 AC AAW61507; standard; peptide; 19 AA.

XX 26-OCT-1998 (first entry)

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 11
 FT /note= "Citruiline"
 XX
 PN MO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 XX
 QY 1 SHOESTRGSRGSGSGS 19
 DB 1 SHOESTRGSRGSGSGS 19
 RESULT 10
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.
 XX
 AC AAW61513;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc9, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruiline"
 FT Modified-site 16
 FT /note= "Citruiline"
 XX
 PN MO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 XX
 QY 1 SHOESTRGSRGSGSGS 19
 DB 1 SHOESTRGSRGSGSGS 19
 RESULT 11
 AAW61505
 ID AAW61505 standard; peptide; 19 AA.
 XX
 AC AAW61505;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc1, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruiline"
 XX
 PN MO9822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 XX
 Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTGRGRGRSGSGS 19
 DB 1 SHQESTYGRGRGRSGSGS 19
 XX
 RESULT 12
 AAM61512
 ID AAM61512 standard; peptide; 19 AA.
 XX
 AC AAM61512;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc8, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT
 FT
 FT
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 PD 14-NOV-1997; 97WO-NL000624.
 XX
 PF 15-NOV-1996; 96NL-01004539.
 XX
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 DR Peptide derived from an antigen recognised by autoantibodies - is
 PT

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 XX
 Query Match 84.8%; Score 76; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.4e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTGRGRGRSGSGS 19
 DB 1 SHQESTYGRGRGRSGSGS 19
 XX
 RESULT 13
 AAM61511
 ID AAM61511 standard; peptide; 19 AA.
 XX
 AC AAM61511;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 11 /note= "Citruilline"
 FT Modified-site 11 /note= "Citruilline"
 FT
 FT
 FT
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 PD 14-NOV-1997; 97WO-NL000624.
 XX
 PF 15-NOV-1996; 96NL-01004539.
 XX
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 DR Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;

Best Local Similarity 84.2%; Pred. No. 1.4e-05; Mismatches 3; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHOESTRGSRGRSGSGS 19
 1 SHOESTRGSRGRSGSGS 19
 DB 1 SHOESTRGSRGRSGSGS 19

RESULT 14

AAW61510
 ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

AC 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citruilline"

FT Modified-site 9 /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.4e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGRSGSGS 19
 1 SHOESTRGSRGRSGSGS 19
 DB 1 SHOESTRGSRGRSGSGS 19

RESULT 15

AAW61520
 ID AAW61520 standard; peptide; 21 AA.

XX AAW61520;

AC 26-OCT-1998 (first entry)

XX Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 4.7e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HOESTRGSRGRSGSGS 19
 4 HOESTRGSRGRSGSGS 21
 DB 4 HOESTRGSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:15
 Job time : 56.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQESTRGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-5
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-3
8	84	91.3	19	9	US-09-308-150-4
9	84	91.3	19	9	US-09-308-150-9
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-7
13	78	84.8	19	9	US-09-308-150-8
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	56	60.9	477	15	US-10-161-927-62	Sequence 62, Appl
17	52	56.5	78	16	US-10-437-963-203073	Sequence 203073,
18	53	53.3	130	16	US-10-437-963-174101	Sequence 174101,
19	49	53.3	1071	16	US-10-437-963-118389	Sequence 118389,
20	48	52.2	122	16	US-10-767-701-32436	Sequence 32436, A
21	48	52.2	172	12	US-10-424-599-215003	Sequence 215003,
22	48	52.2	188	12	US-10-425-114-39730	Sequence 39730, A
23	48	52.2	211	14	US-10-424-599-157696	Sequence 157696,
24	48	52.2	211	14	US-10-002-631C-268	Sequence 268, App
25	47	51.1	184	15	US-10-369-493-6474	Sequence 6474, Ap
26	47	51.1	372	9	US-09-925-300-1619	Sequence 1619, Ap
27	47	51.1	372	16	US-10-437-963-143403	Sequence 143403,
28	47	51.1	436	14	US-10-156-761-13022	Sequence 13022, A
29	47	51.1	492	16	US-10-437-963-120908	Sequence 120908,
30	47	51.1	511	16	US-10-437-963-188040	Sequence 188040,
31	47	51.1	563	16	US-10-437-963-198539	Sequence 198539,
32	47	51.1	838	16	US-10-437-963-146503	Sequence 146503,
33	47	51.1	854	12	US-10-425-114-57838	Sequence 57838, A
34	46	50.0	16	10	US-09-791-524-150	Sequence 150, App
35	46	50.0	142	16	US-10-767-701-49555	Sequence 49555, A
36	46	50.0	178	16	US-10-437-963-142959	Sequence 142959,
37	46	50.0	231	16	US-10-437-963-179346	Sequence 179346,
38	46	50.0	345	12	US-10-425-114-71489	Sequence 71489, A
39	46	50.0	461	16	US-10-437-963-185969	Sequence 185969,
40	46	50.0	1087	9	US-09-918-909-24	Sequence 918, Appl
41	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
42	45	48.9	49	12	US-10-424-599-187406	Sequence 187406,
43	45	48.9	150	16	US-10-767-701-52043	Sequence 52043, A
44	45	48.9	181	16	US-10-437-963-164001	Sequence 164001,
45	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A

ALIGNMENTS

RESULT 1
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoelt, Rene Michael Antonius
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 2

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19
 ; TYPE: PRP
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 3

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 19

TYPE: PRP
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRP
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 5

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```

; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 6
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 7
US-09-308-150-3

```

```

; Sequence 3, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

```

```

RESULT 8
US-09-308-150-4
; Sequence 4, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline

```

US-09-308-150-4

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 9
US-09-308-150-9

; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 3e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 10
US-09-308-150-12

; Sequence 12, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624

; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539

; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 11
US-09-308-150-6

; Sequence 6, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00021;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 12
US-09-308-150-7

; Sequence 7, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius

Query Match	84.8%	Score 78;	DB 9;	Length 19;
Best Local Similarity	84.2%	Pred: No.	0.00021;	
Matches	16;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;

RESULT 13
US-09-308-150-8
: Sequence 8, Application US/09308150
: Patent No. US20020137092A1

```

APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoelt, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIORITY FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ. ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO. 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-B
```

Query March	84.8%	Score 78;	DB 9;	Length 19;
Best Local Similarity	84.2%	Pred. 0.00027;		
Matches 16; Conservative	0;	Mismatches 13;	Indels 0;	Gaps 0;

RESULT 14
US-09-308-150-10

```

US-09-308-150-10
/ Sequence 10, Application US/09308150
/ Patent No. US20020.37092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Schelkundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ. ID NOS.: 16
/ SOFTWARE: Patencin ver. 2.1
/ SEQ. ID NO. 10
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: known CDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
/ NAME/KEY: DISULFID
/ LOCATION: (3). (16)
/ IS-09-308-150-10

```

Query Match	81.5%	Score 75	DB 9	Length 21
Best Local Similarity	83.3%	Pred. No.	0.00063	
Matches 15, Conservative		Mismatches 3	Indels 0	Gaps 0

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1

```

: APPLICANT: Moereels, Henri
: APPLICANT: Mebeus, Lydie
: TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS
: FILE REFERENCE: 11362.003INPUS00 INNS:031
: CURRENT APPLICATION NUMBER: US/09/747,029A
: CURRENT FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: EP 00870195.5
: PRIOR FILING DATE: 2000-09-08
: PRIOR APPLICATION NUMBER: EP 99870280.7
: PRIOR FILING DATE: 1999-12-21
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 22
: LENGTH: 22
: TYPE: PRT

```


Tue Sep 28 07:11:33 2004

us-09-308-150-5.open.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-5

Perfect score: 92

Sequence: 1 SHQESTGRSGRSGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
1	47	51.1	629	4	US-09-252-991A-22901
2	44	47.8	157	4	US-09-252-991A-27836
3	44	47.8	653	4	US-09-513-057C-13
4	43	46.7	159	4	US-09-252-991A-27107
5	43	46.7	581	4	US-09-252-991A-25519
6	43	46.7	722	4	US-09-984-890-4
7	43	46.7	724	4	US-09-984-890-2
8	43	46.7	745	4	US-09-523-849-36
9	43	46.7	2237	1	US-08-354-973-1
10	42.5	46.2	398	4	US-09-252-991A-19488
11	42	45.7	105	4	US-09-252-991A-32394
12	42	45.7	349	4	US-09-252-991A-27336
13	42	45.7	523	4	US-09-252-991A-21124
14	41	44.6	70	2	US-08-724-548-54
15	41	44.6	70	3	US-07-978-6748-54
16	41	44.6	154	4	US-09-252-991A-32846
17	41	44.6	169	3	US-09-342-084-6
18	41	44.6	204	4	US-09-252-991A-21317
19	41	44.6	205	4	US-09-252-991A-18529
20	41	44.6	255	4	US-09-252-991A-19598
21	41	44.6	357	4	US-09-252-991A-17216
22	41	44.6	409	4	US-09-489-039A-14217
23	41	44.6	448	4	US-09-252-991A-24066
24	41	44.6	482	4	US-09-252-991A-16654
25	41	44.6	719	4	US-09-252-991A-28576
26	41	44.6	722	4	US-08-817-832B-32
27	40.5	44.0	257	4	US-09-252-991A-22204

28	40.5	44.0	703	3	US-08-910-925-4	Sequence 4, Appli
29	40	43.5	134	4	US-09-252-991A-18314	Sequence 18314, A
30	40	43.5	170	4	US-09-252-991A-26306	Sequence 26306, A
31	40	43.5	237	4	US-09-252-991A-23304	Sequence 23304, A
32	40	43.5	239	4	US-09-252-991A-28985	Sequence 28985, A
33	40	43.5	255	3	US-09-342-084-11	Sequence 11, Appl
34	40	43.5	274	4	US-09-252-991A-16696	Sequence 16696, A
35	40	43.5	286	4	US-09-252-991A-22717	Sequence 22717, A
36	40	43.5	294	4	US-09-252-991A-21947	Sequence 21947, A
37	40	43.5	326	4	US-09-252-991A-20075	Sequence 20075, A
38	40	43.5	363	4	US-09-252-991A-25132	Sequence 25132, A
39	40	43.5	406	4	US-09-252-991A-19857	Sequence 19857, A
40	40	43.5	421	4	US-09-252-991A-32326	Sequence 32326, A
41	40	43.5	478	4	US-09-252-991A-25411	Sequence 25411, A
42	40	43.5	480	4	US-09-252-991A-22879	Sequence 22879, A
43	40	43.5	614	4	US-09-252-991A-29695	Sequence 29695, A
44	40	43.5	676	2	US-08-398-590A-40	Sequence 40, Appl
45	40	43.5	676	3	US-08-894-997-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-22901
; Sequence 22901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22901
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22901

Query Match 51.1%; Score 47; DB 4; Length 629;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 QESTGRSGRSGRSGS 19
Db 272 RQSHGRAGRAGRSGS 288

RESULT 2
US-09-252-991A-27836
; Sequence 27836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27836
; LENGTH: 157
; TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27836
Query Match      47.8%; Score 44; DB 4; Length 157;
Best Local Similarity 66.7%; Pred. No. 9.5;
Matches      8; Conservative      2; Mismatches      2; Indels      0; Gaps      0;

QY      6 TGGSRGRSGXS 17
      :|||||:
Db      70 SRGRGRSGSA 81

RESULT 3
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
; US-09-513-057C-13

Query Match      47.8%; Score 44; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 42;
Matches      10; Conservative      1; Mismatches      8; Indels      0; Gaps      0;

QY      1 SHQSTRGRSGRSGXS 19
      :|||||:
Db      573 SRQVTSASASGREGISGS 591

RESULT 4
US-09-252-991A-27107
; Sequence 27107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27107
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27107

Query Match      46.7%; Score 43; DB 4; Length 159;
Best Local Similarity 57.9%; Pred. No. 14;
Matches      11; Conservative      0; Mismatches      8; Indels      0; Gaps      0;

QY      1 SHQSTRGRSGRSGXS 19
      :|||||:
Db      5 SASSTRGRSAMRSTASS 23

RESULT 5
US-09-252-991A-25519
; Sequence 25519, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25519
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25519

Query Match      46.7%; Score 43; DB 4; Length 581;
Best Local Similarity 50.0%; Pred. No. 53;
Matches      9; Conservative      2; Mismatches      7; Indels      0; Gaps      0;

QY      2 HQSTRGRSGRSGXS 19
      :|||||:
Db      178 HRADSRGRSRERGAPRS 195

RESULT 6
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-984-890-4

Query Match      46.7%; Score 43; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 67;
Matches      9; Conservative      1; Mismatches      5; Indels      0; Gaps      0;

QY      5 STRGRSGRSGXS 19
      :|||||:
Db      562 SPGHSGRGRGASGS 576

RESULT 7
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-890-2

Query Match
Best Local Similarity 46.7%; Score 43; DB 4; Length 724;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGSRGRSGXSGS 19
DB 564 SPGSHSQRRGASGS 578

RESULT 8
US-09-523-849-36
; Sequence 36; Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match
Best Local Similarity 46.7%; Score 43; DB 4; Length 745;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGSRGRSGXSGS 19
DB 585 SPGSHSQRRGASGS 599

RESULT 9
US-08-354-973-1
; Sequence 1; Application US/08354973
; Patent No. 5641666
; GENERAL INFORMATION:
; APPLICANT: Vanlensieck, Hans-Friedrich
; APPLICANT: Himmen, Albert
; TITLE OF INVENTION: Fungi Resistant to Soraphen A
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: Patent Dept., 520 White Plains Rd., POB 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/354,973
; FILING DATE: 13-Dec-1994
```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: PF/5-19802/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHEICAL: NO
US-08-354-973-1

Query Match
Best Local Similarity 46.7%; Score 43; DB 1; Length 2237;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSRGRSGXSGS 19
DB 1200 HOSSNGPAPDRSGSSAS 1217

RESULT 10
US-09-252-991A-19488
; Sequence 19488; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19488
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19488

Query Match
Best Local Similarity 46.2%; Score 42.5; DB 4; Length 398;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 HOESTRGRS-RGRSG 15
DB 326 HRSRRGRSGRGRG 340

RESULT 11
US-09-252-991A-32394
; Sequence 32394; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32394
LENGTH: 105
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32394

Query Match 45.7%; Score 42; DB 4; Length 105;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 TRGSRGRSGXSG 18
DB 65 TRGRCACRCATG 77

RESULT 12
US-09-252-991A-27336
Sequence 27336, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27336
LENGTH: 349
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27336

Query Match 45.7%; Score 42; DB 4; Length 349;
Best Local Similarity 52.6%; Pred. No. 45;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGXSGS 19
DB 309 SRPASGRGRCRCSAGAS 327

RESULT 13
US-09-252-991A-21124
Sequence 21124, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21124
LENGTH: 523
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21124

Query Match 45.7%; Score 42; DB 4; Length 523;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGXSG 18
DB 85 SRPAATRGRRRRRGLEG 102

RESULT 14
US-08-724-548-54
Sequence 54, Application US/08724548
Patent No. 5830637
GENERAL INFORMATION:
APPLICANT: Frank, Ronald
APPLICANT: Guler, Sinan
TITLE OF INVENTION: PROCESS FOR THE RAPID SYNTHESIS OF
TITLE OF INVENTION: SUPPORT-BOUND OR FREE PEPTIDES OR
TITLE OF INVENTION: OLIGONUCLEOTIDES, A FLAT MATERIAL,
TITLE OF INVENTION: PREPARED THEREWITH AND ITS USE AND
TITLE OF INVENTION: A DEVICE FOR CARRYING OUT THE
PROCESS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSER: Joseph T. Eisele
ADDRESSER: Kane, Dalsimer, Sullivan, Kurucz,
ADDRESSER: Levy, Eisele and Richard
STREET: 711 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017-4059
COMPUTER READABLE FORM:
MEDIUM TYPE: 3-1/2" DISKETTE
COMPUTER: IBM-XT COMPATIBLE
OPERATING SYSTEM: DOS 3.3:
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,548
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,674
FILING DATE: 03/22/93
ATTORNEY/AGENT INFORMATION:
NAME: EISELE, JOSEPH T.
REGISTRATION NUMBER: 25,331
REFERENCE/DOCKET NUMBER: 2727-68 PCT
TELEPHONE: (212) 687-6000
TELEFAX: (212) 682-3485
TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 residues
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: Peptide
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HABLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
US-08-724-548-54

Query Match 44.6%; Score 41; DB 2; Length 70;
Best Local Similarity 61.5%; Pred. No. 12;

Tue Sep 28 07:11:33 2004

us-09-308-150-5.open.ra1

Page 5

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 7 RGRSRGRSGXSGS 19
Db 7 RGRSRGRSGXSGS 19

RESULT 15
US-07-978-674B-54
; Sequence 54, Application US/07978674B
; Patent No. 6040423
; GENERAL INFORMATION:

APPLICANT: Frank, Ronald
TITLE OF INVENTION: Guler, Sinan
TITLE OF INVENTION: PROCESS FOR THE RAPID SYNTHESIS OF
TITLE OF INVENTION: SUPPORT-BOUND OR FREE PEPTIDES OR
TITLE OF INVENTION: OLIGONUCLEOTIDES, A FLAT MATERIAL
TITLE OF INVENTION: PREPARED THERMITH AND ITS USE AND
TITLE OF INVENTION: A DEVICE FOR CARRYING OUT THE
TITLE OF INVENTION: PROCESS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Joseph T. Eisele
ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,
ADDRESSEE: Levy, Eisele and Richard
STREET: 711 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017-4059
COMPUTER READABLE FORM:
MEDIUM TYPE: 3-1/2" DISKETTE
COMPUTER: IBM-XT COMPATIBLE
OPERATING SYSTEM: DOS 3.3:
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/978,674B
FILING DATE: 03/22/93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: EISELE, JOSEPH T.
REGISTRATION NUMBER: 25,331
REFERENCE/DOCKET NUMBER: 2727-68 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 687-6000
TELEFAX: (212) 682-3485
TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 residues
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: Peptide
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
US-07-978-674B-54

Query Match 44.6%; Score 41; DB 3; Length 70;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 7 RGRSRGRSGXSGS 19
Db 7 RGRSRGRSGXSGS 19

Search completed: September 28, 2004, 06:26:43
Job time : 14.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-6
Perfect score: 88
Sequence: 1 SHOESTXGXRGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	2248	2	A35938
2	78	88.6	416	2	A32947
3	69	78.4	591	2	A45135
4	52	59.1	506	1	W2ML47
5	49	55.7	822	2	T51049
6	47	53.4	399	2	T46259
7	44	50.0	800	2	T02852
8	44	50.0	827	2	T13468
9	44	50.0	842	1	JDVLVS
10	44	50.0	843	1	JDVLVR
11	43	48.9	306	2	T21220
12	43	48.9	694	2	S71786
13	43	48.9	849	2	A86592
14	42	47.7	136	2	T35632
15	42	47.7	315	2	S53589
16	42	47.7	471	2	T33997
17	42	47.7	745	2	G01025
18	42	47.7	774	2	I48609
19	42	47.7	889	2	T33422
20	42	47.7	1337	2	T30291
21	42	47.7	1804	2	T34518
22	41.5	47.2	1895	2	T06609
23	41	46.6	825	2	GC1163
24	41	46.6	836	2	G84727
25	41	46.6	843	1	T43491
26	41	46.6	1015	2	T42013
27	41	46.6	2233	2	S63347
28	40	45.5	163	2	T23076
29	40	45.5	167	2	S05545

30	40	45.5	168	2	A39316	RAB-17 protein - m
31	40	45.5	168	2	S08633	RAB-17 protein - m
32	40	45.5	197	2	T18918	hypothetical prote
33	40	45.5	203	2	C87801	proteih C10G11.9 l
34	40	45.5	261	2	S33356	hypothetical prote
35	40	45.5	269	2	B38095	T-cell-specific tr
36	40	45.5	272	2	A38900	T-cell-specific tr
37	40	45.5	303	2	JH0401	T-cell-specific tr
38	40	45.5	312	2	A31846	130K paracrystall
39	40	45.5	508	2	T22440	hypothetical prote
40	40	45.5	538	2	H65335	T20H2.2 protein -
41	40	45.5	611	2	T06458	nucleolin homolog
42	40	45.5	649	2	G71283	probable ATP-depen
43	40	45.5	692	2	H71426	hypothetical prote
44	40	45.5	695	2	G85135	hypothetical prote
45	40	45.5	708	2	T25725	hypothetical prote

ALIGNMENTS

RESULT 1

A35938
Profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence__revision 02-Jul-1996 #text__change 29-Sep-1999
C/Accession: A35938
R/Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinhert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; PMID:91064347; PMID:2248557
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:J02929
C/Genetics:
A/Gene: GDB:RUG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: profilaggrin repeat
F:570-893/Region: profilaggrin repeat
F:1074-1397/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 95.5%; Score 84; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 551 SHOESTXGXRGRSGRSGS 569

RESULT 2

A32947
Filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence__revision 04-Sep-1992 #text__change 29-Sep-1999
R/McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of th
A/Reference number: A32947; PMID:8926901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-references: GB:M2435; NID:g182604; PTDN:AA52454.1; PID:g182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

A;Gene: GDB:FLG
A;Cross-references: GDB:119912; OMIM:135940
A;Map position: 1q21-1q21
C;Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C;Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 88.6%; Score 78; DB 2; Length 416;
Best Local Similarity 84.2%; Pred. No. 4.9e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGSRGSGS 19
DB 7 SHQESTRGRSAGSRGSGS 25

RESULT 3
A45135
Profilaggrin - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C;Accession: A45135
R;Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nurusukirti, W.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992
A;Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A;Reference number: A45135; MUID:93054736; PMID:1429717
A;Accession: A45135
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-591 <PR>
A;Cross-references: GB:L01089; GB:M90967; NID:G190408; PID:AAA60177.1; PID:9553621
A;Note: sequence extracted from NCBI backbone (NCBIP:118773)
C;Genetics:
A;Gene: GDB:FLG
A;Cross-references: GDB:119912; OMIM:135940
A;Map position: 1q21-1q21
C;Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C;Keywords: EF hand; epidermis; polymorphism; tandem repeat
F;49-81/Domain: calmodulin repeat homology <EP2>

Query Match 78.4%; Score 69; DB 2; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.0019;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGSRGSGS 19
DB 449 SHQESTRGRSAGSRGSGS 467

RESULT 4
W2W147
E2 protein - human papillomavirus type 47
C;Species: human papillomavirus type 47
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C;Accession: D35324
R;Kiyono, T.; Adachi, A.; Ishibashi, M.
Virology 177, 401-405, 1990
A;Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
A;Reference number: A35324; MUID:90281611; PMID:2162112
A;Accession: D35324
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-506 <KIT>
A;Cross-references: GB:M23305; NID:G333062; PID:AAA46979.1; PID:G333067
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

Query Match 59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.89;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGSRGSGS 19
DB 1 SHQESTXGSRGSRGSGS 19

DB 342 SRQNTNRGRGRGRGSGS 360

RESULT 5
T51049
related to nucleolar phosphoprotein [imported] - Neurospora crassa
N;Alternate names: protein B12P1.10
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C;Accession: T51049
R;Schulte, U.; Algn, V.; Hobeisel, U.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25286
A;Accession: T51049
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-822 <SCH>
A;Cross-references: EMBL:AL390091; GSPDB:GNO0116; NCSP:B12P1.10
A;Experimental source: BAC clone B12P1, strain OR74A
C;Genetics:
A;Gene: NCSP:B12P1.10
A;Map position: 6
A;Intons: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;
Best Local Similarity 52.9%; Pred. No. 4.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 414 RESASGRTRGRGRGSGT 430

RESULT 6
T46259
hypothetical protein DKFZp761E0323.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T46259
R;Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23032
A;Accession: T46259
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-399 <AAA>
A;Cross-references: EMBL:AL137423
A;Experimental source: adult amygdala; clone DKFZp761E0323
C;Genetics:
A;Note: DKFZp761E0323.1

Query Match 53.4%; Score 47; DB 2; Length 399;
Best Local Similarity 55.6%; Pred. No. 4.5;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTXGSRGSRGSGS 18
DB 336 SNERQRSGWSGSRGSGS 353

RESULT 7
T02852
probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)
C;Species: Leishmania major
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C;Accession: T02852; H81462
R;Wyley, P.J.
submitted to the EMBL Data Library, May 1998
A;Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
A;Reference number: Z14740
A;Accession: T02852
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A/Residues: 1-800 <MTL>
 A/Cross-references: EMBL:AE001274; NID:G3264850; PID:G2266920
 R/Miyler, P.J.; Audleman, L.J. deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A/Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A/Reference number: AB1455; MUID:99178987; PMID:10077609
 A/Accession: AB1462
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-800 <PYL>
 A/Cross-references: GB:AE001274; NID:G3264850; PID:AC24675.1; PID:G2266920; GSPDB:GN00
 A/Experimental source: strain MHW/IL/81/Friedlin
 C/Genetics:
 A/Gene: L1439.4
 A/Map position: 1

Query Match 50.0%; Score 44; DB 2; Length 800;
 Best Local Similarity 50.0%; Pred. No. 27;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 HQSTXGSRGRSGSGS 19
 ||:|||||
 Db 429 HRDGVRLSTVSGSGS 446

RESULT 8
 T13468
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)
 C/Species: hepatitis B virus, HBV
 A/Variety: isolate 09D09HCC
 C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
 C/Accession: T13468
 R/Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishihiro, S.
 Arch. Virol. 143, 2313-2326, 1998
 A/Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcino
 A/Reference number: Z17684; MUID:99129050; PMID:9930189
 A/Accession: T13468
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-827 <TK>
 A/Cross-references: EMBL:AB014368; NID:G3551304; PID:BA32863.2; PID:G6116707
 A/Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
 C/Genetics:
 A/Gene: P
 A/Introns: 303/3
 C/Superfamily: hepatitis virus DNA-directed DNA polymerase
 C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 2; Length 827;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILARKSGSGS 242

RESULT 9
 JDUVLR
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
 C/Species: hepatitis B virus, HBV
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 25-Oct-1996
 C/Accession: S04568
 R/Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
 Nucleic Acids Res. 17, 2124, 1989
 A/Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp
 A/Reference number: S04568; MUID:89183619; PMID:2928116
 A/Accession: S04568
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-842 <RHO>
 A/Cross-references: EMBL:X14193
 C/Superfamily: hepatitis virus DNA-directed DNA polymerase

C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 1; Length 842;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILARKSGSGS 242

RESULT 10
 JDUVLR
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
 C/Species: hepatitis B virus, HBV
 C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 25-Oct-1996
 C/Accession: A00704
 R/Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
 Nucleic Acids Res. 11, 1747-1757, 1983
 A/Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
 A/Reference number: A93460; MUID:83168919; PMID:6300776
 A/Accession: A00704
 A/Molecule type: DNA
 A/Residues: 1-843 <ONO>
 C/Superfamily: hepatitis virus DNA-directed DNA polymerase
 C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 1; Length 843;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILARKSGSGS 242

RESULT 11
 T21220
 hypothetical protein F21H7.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C/Accession: T21220
 R/Gardner, A.
 submitted to the EMBL Data Library, March 1997
 A/Reference number: Z19393
 A/Accession: T21220
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-306 <MTL>
 A/Cross-references: EMBL:Z93379; PID:CA807588.1; GSPDB:GN00023; CESP:F21H7.5
 A/Experimental source: clone F21H7
 C/Genetics:
 A/Gene: CESP:F21H7.5
 A/Map position: 5
 A/Introns: 28/3; 146/3; 232/3
 C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHQSTXGSRGRSGSGS 18
 ||:|||||
 Db 48 SHRSKTKRSGSGSGS 65

RESULT 12
 S71786
 wingless receptor precursor dfw2 - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 C/Accession: S71786; S78444
 R/Bhanot, P.; Brink, M.; Harryman Samos, C.; Heish, J.C.; Wang, Y.; Macke, J.P.; Andrew,

Nature 382, 225-230, 1996
 A>Title: A new member of the frizzled family from *Drosophila* functions as a wingless rec
 A/Reference number: S71786; PMID:96353971; PMID:8717036
 A/Accession: S71786
 A>Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1694 <BHA>
 A/Cross-references: EMBL:U65589
 A/Note: mRNA was also sequenced
 R/Bhanot, P.; Wang, Y.; Nathans, J.
 submitted to the EMBL Data Library, July 1996
 A/Reference number: S78444
 A/Accession: S78444
 A/Molecule type: DNA
 A/Residues: 1416, 418-694 <BHM>
 A/Cross-references: EMBL:U65589; NID:G1518050; PIDN:AAC47273.1; PID:G1518051
 A/Genetics: C
 A/Cross-references: FlyBase:FBgn0016797
 C/Superfamily: fruit fly frizzled protein
 C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 Oy 1 SHOESTXGSGRSGRSGS 19
 Db 182 SYTBAGSGSGSGSGSGS 200

RESULT 13
 A96592
 hypothetical protein F14C21.55 [imported] - *Arabidopsis thaliana*
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 A/Accession: A96592
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Hansen, N.F.; Hughes, B.; Hultzar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Matzli, H.
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A/Reference number: A86141; PMID:21016719; PMID:11130712
 A/Accession: A96592
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-849 <STO>
 A/Cross-references: GB:AE005173; NID:G11055759; PIDN:AG28231.1; GSPDB:GN00141
 C/Genetics: C
 A/Genetics: F14C21.55
 A/Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 5 STXGSGRSGRSGS 19
 Db 736 SSANRSGRSGRSGS 750

RESULT 14
 T35632
 probable transposase - *Streptomyces coelicolor*
 C/Species: *Streptomyces coelicolor*
 C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
 A/Accession: T35632
 R/Seeger, K.U.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999
 A/Reference number: 221584
 A/Accession: T35632
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-136 <SEE>
 A/Cross-references: EMBL:AU079356; PIDN:CAB45627.1; GSPDB:GN00070; SCOEDB:SCG69.36C
 A/Experimental source: strain A3(2)
 C/Genetics: C
 A/Genetics: SCOEDB:SCG69.36C
 C/Superfamily: *Streptomyces coelicolor* probable transposase SCG69.36C

Query Match 47.7%; Score 42; DB 2; Length 136;
 Best Local Similarity 41.2%; Pred. No. 9.8;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 SHOESTXGSGRSGRSGS 17
 Db 111 AHQHAAGARRKRGRRGRA 127

RESULT 15
 S53589
 SOL2 protein - yeast (*Saccharomyces cerevisiae*)
 N/Alternate names: protein YCR073W-a; protein YCRX13W
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 19-May-1995 #sequence_revision 17-Nov-1995 #text_change 19-Apr-2002
 A/Accession: S53589; S70385; S53593
 R/Jimenez, A.
 submitted to the EMBL Data Library, December 1992
 A/Reference number: S53589
 A/Accession: S53589
 A/Molecule type: DNA
 A/Residues: 1-315 <JIM>
 A/Cross-references: EMBL:X59720; NID:G1907116; PID:G1907213
 R/Shen, W.C.; Stanford, D.R.; Hopper, A.K.
 Genetics 143, 699-712, 1996
 A/Title: *Los1p*, involved in yeast pre-tRNA splicing, positively regulates members of the
 A/Reference number: S70383; PMID:96363906; PMID:8725220
 A/Accession: S70385
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-170, 172-315 <SHE>
 A/Cross-references: EMBL:U46559; NID:G1184940; PIDN:AA849322.1; PID:G1184941
 C/Genetics: C
 A/Genetics: SGD:SOL2; SOL2
 A/Cross-references: MIPS:YCR073W-a; SGD:S0000718
 A/Map position: 3R
 C/Function: multicopy suppressor of *los1-1*

Query Match 47.7%; Score 42; DB 2; Length 315;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Oy 4 ESTXGSGRSGRSGS 19
 Db 60 KSTASABKSGSGSGS 75

Search completed: September 28, 2004, 06:15:01
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds

(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-6

Perfect score: 88
Sequence: 1 SHQSTXGXSRRGSRGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	88.6	416	1 FILA_HUMAN	P20930 homo sapien
2	52	59.1	506	1 VEB_HPV47	P22420 human papil
3	44	50.0	730	1 DPOL_HPBV4	P12933 hepatitis b
4	44	50.0	842	1 DPOL_HPBV4	P13870 hepatitis b
5	44	50.0	843	1 DPOL_HPBV4	P03157 hepatitis b
6	43	48.9	694	1 FRZ2_DROME	P39VX3 dirosophila
7	42	47.7	315	1 SOL2_YEAST	P37262 saccharomyc
8	42	47.7	774	1 MRX2_MOUSE	O05512 mus musculu
9	42	47.7	1337	1 DEXT_STRDO	P39653 streptococc
10	41.5	47.2	1895	1 WR19_ARATH	O9867 arabidopsis
11	41	46.6	825	1 SRS_RAT	O63003 rattus norv
12	41	46.6	1015	1 FRC_CRESP	O00586 creopus spi
13	41	46.6	2233	1 COX6_YEAST	O00855 saccharomyc
14	40.5	46.0	379	1 ROA3_MOUSE	O80905 mus musculu
15	40	45.5	168	1 DH1_MAZE	P12950 zea mays (m
16	40	45.5	380	1 PEXD_PICPA	O92266 pichia past
17	40	45.5	419	1 TCF7_MOUSE	O00417 mus musculu
18	40	45.5	660	1 DDX7_MOUSE	O15523 mus musculu
19	40	45.5	778	1 SEA2_MYCLE	O32822 mycobacteri
20	40	45.5	808	1 SEA2_MYCTU	O50612 mycobacteri
21	40	45.5	955	1 CLS2_HUMAN	O98400 homo sapien
22	39.5	44.9	135	1 RBP1_DROME	O04227 dirosophila
23	39	44.3	155	1 RRP7_PSIINU	O88WV0 psilocybe nu
24	39	44.3	439	1 AP5G_DICDI	O54672 dictyosteli
25	39	44.3	508	1 Y202_HUMAN	O92599 homo sapien
26	39	44.3	686	1 BDR1_YEAST	P35617 saccharomyc
27	38	43.2	295	1 Y42A_RHIST	P55728 rhizobium s
28	38	43.2	373	1 YF04_YEAST	P43603 saccharomyc
29	38	43.2	502	1 VEB_HPV25	P36787 human papil
30	38	43.2	643	1 GAG_SFPV3L	P27400 simian foam
31	38	43.2	645	1 R1P4_MOUSE	O8161 mus musculu
32	38	43.2	651	1 SEC9_YEAST	P40357 saccharomyc
33	38	43.2	824	1 TGM1_RAT	P23606 rattus norv

34	38	43.2	846	1 IF2_PSEPK	O884V7 pseudomonas
35	38	43.2	965	1 CLS2_RAT	O8Vd41 rattus norv
36	38	43.2	966	1 CLS2_MOUSE	O9865 mus musculu
37	38	43.2	978	1 PEX6_RAT	P54777 rattus norv
38	38	43.2	980	1 PEX6_HUMAN	O13608 homo sapien
39	38	43.2	1094	1 A3B1_HUMAN	O00203 homo sapien
40	38	43.2	1232	1 Y005_CAEEL	P34643 caenorhabd1
41	38	43.2	1383	1 PRAX_RAT	O63425 rattus norv
42	38	43.2	1733	1 VNVA_PRIVKA	P33485 pseudorabie
43	37.5	42.6	832	1 DPOL_HPBV4	P12900 hepatitis b
44	37	42.0	123	1 LSM4_CAEEL	O19522 caenorhabd1
45	37	42.0	136	1 SRI9_ORYSA	P49964 oryza sativ

ALIGNMENTS

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
ID FILA_HUMAN
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69296901; PubMed=2740331;
RA McKinley-Grant U.U., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RP CITRULLINATION.
RX MEDLINE=96374388; PubMed=8780679;
RA Senshu T., Kan S., Ogawa H., Manabe M., Aaga H.;
RT "Preferential determination of keratin Ki and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -!- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -!- PTM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -!- PTM: Undergoes deamination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M24355; AAA52454.1; -
DR PIR: A32947; A32947.
DR Genew; HGNC:3748; FLG.
DR MIM: 135940; -
DR GO: GO:0005882; C:intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007275; P:development; NAS.
DR InterPro: IPR003303; Filaggrin.


```

RP SEQUENCE FROM N.A.
RX MEDLINE=89183619; PubMed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
RT virus subtype adr.";
RL Nucleic Acids Res. 17:2124-2124(1989).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonomester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14193; CAA32399.1; ALT_TERM.
DR EMBL; X14193; CAA32405.1; ALT_TERM.
DR PIR; S04568; JDVLVS.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvt; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
KM Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KM Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 842 AA; 94545 MW; 137FE37246BA5CC6 CRC64;

Query Match 50.0%; Score 44; DB 1; Length 842;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGXSGRSGRSGS 19
Db 226 QPQGSWARGSGRSGS 242

RESULT 5
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID DPOL_HPBVR
AC F03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=106820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adw.";
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonomester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00867; NOT_ANNOTATED_CDS.
DR PIR; A00704; JDVLVR.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvt; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
KM Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KM Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D49083C4E8B CRC64;

Query Match 50.0%; Score 44; DB 1; Length 843;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGXSGRSGRSGS 19
Db 226 QPQGSWARGSGRSGS 242

RESULT 6
FRZ2_DROME STANDARD; PRT; 694 AA.
ID FRZ2_DROME
AC Q9VFX3; Q94916; Q9VFX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Frizzled protein 2 precursor (Frizzled-2) (dfrz2).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A. AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
RX MEDLINE=96353971; PubMed=8717036;
RA Bhano P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
RA Andrew D., Nathans J., Nuse R.;
RT "A new member of the frizzled family from Drosophila functions as a
RT Wingless receptor";
RL Nature 382:225-230(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Berens P.V., Bernan B.P., Bhandari D., Bolashkov S.,
RA Borokova D., Botchan M.R., Bouck J., Broksrein P., Brothier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacted J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins. Inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeleton of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage
 CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC
 CC -1- DOMAIN: Lys-Thr-X-X-TRP motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC
 CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: U65589; AAC47273.1; -;
 CC EMBL: AE003518; AAF49185.2; -;
 CC PIR: S71786.
 CC RYBase: FBgn0016797; fz2.
 CC
 CC GO: GO:016021; C: integral to membrane; NMS.
 CC GO: GO:001747; F: Wnt-protein binding; IDA.
 CC GO: GO:0006928; P: cell motility; IMP.
 CC GO: GO:0007163; P: establishment and/or maintenance of cell po. . .; ISS.
 CC GO: GO:0005855; P: female gonad development; IMP.
 CC GO: GO:0016055; P: Wnt receptor signaling pathway; IDA.
 CC InterPro: IPR000539; Frizzled.
 CC InterPro: IPR000034; Fz domain.
 CC InterPro: IPR000832; GPCR_sectretin.
 CC Pfam: PF01534; Frizzled_1.
 CC Pfam: PF01392; Fz_1.
 CC PRINTS: PRO0489; FRIZZLED.
 CC SMART: SM00063; FRI; 1.
 CC PROSITE: PSS0038; FZ; 1.
 CC PROSITE: PSS0261; G_PROTEIN_RECEP_F2_4; 1.

KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
 KM Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 694
 FT DOMAIN 23 315
 FT DOMAIN 316 336
 FT DOMAIN 337 352
 FT DOMAIN 353 373
 FT DOMAIN 374 397
 FT DOMAIN 398 418
 FT DOMAIN 419 439
 FT DOMAIN 440 460
 FT DOMAIN 461 482
 FT DOMAIN 483 503
 FT DOMAIN 504 534
 FT DOMAIN 535 555
 FT DOMAIN 556 584
 FT DOMAIN 585 605
 FT DOMAIN 606 694
 FT DOMAIN 187 225
 FT DOMAIN 59 180
 FT SITE 608 613
 FT SITE 692 694
 FT CARBOHYD 78 78
 FT CARBOHYD 288 288
 FT CONFLICT 55 55
 FT CONFLICT 417 417
 SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAFB096 CRC64;
 Query March 48.9%; Score 43; DB 1; Length 694;
 Best Local Similarity 52.6%; Pred. No. 14;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHOSTYXSGRGRSGSGS 19
 DB 182 SYTAGSGSGSGSGSGS 200
 RESULT 7
 SOL2_YEAST STANDARD; PRT; 315 AA.
 AC P37262;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (PGLL).
 GN SOL2 OR YCRX13W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96363906; PubMed=8725220;
 RA Shen W.C., Stanford D.R., Hopper A.K.;
 RT "Lcsp1, involved in yeast pre-tRNA splicing, positively regulates
 RT members of the SOL gene family.";
 RL Genetics 143:699-712(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ballesca J.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
 RA Sanz E.;
 RT Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION, AND SIMILARITY.
 RX MEDLINE=94147996; PubMed=8313894;
 RA Kocin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RL EMBO J. 13:493-503(1994).
 CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
 CC phosphogluconate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
 CC phospho-D-gluconate.

```

CC -1- PATHWAY: Pentose phosphate pathway; second step.
CC -1- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate
CC isomerase family. 6-phosphoglucosaminolactonase subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U46559; AA349322.1; -.
DR EMBL: X59720; CAA42272.1; -.
DR PIR: S53589; S53589.
DR Germonline: 139004; -.
DR SGD: S0000718; SOL2.
DR GO: GO:0008033; P:RNA processing; IGI.
DR InterPro: IPR006148; Gluc gal isom.
DR InterPro: IPR005900; Phosphoglucolac.
DR Pfam: PF01182; Glucosamine iso; 1.
DR TIGRFAMs: TIGR01198; pgl1.1.
KM Hydrolyase.
FT CONFLICT 171 171 A -> V (IN REF. 2).
SQ SEQUENCE 315 AA; 34501 MW; 7F0AAD76574AB276 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 315;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 ESTXGSRGSRGSGS 19
Db : : : : :
60 KSTASAAEGKSGSGS 75

RESULT 9
MRK2_MOUSE STANDARD; PRT; 774 AA.
ID MRK2_MOUSE STANDARD; PRT; 774 AA.
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
GN MARK2 OR EMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93364122; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19.";
RL Mamm. Genome 4:401-403 (1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X70764; CAA50040.1; -.
DR PIR: I48609; I48609.
DR HSSP: Q63450; IAO6.
DR MGD; MGI:99636; Marfk2.

```

```

DR InterPro: IPR001772; Kinase Cterm.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF02149; KAI; 1.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 53 304
FT NP_BIND 59 67 ATP (BY SIMILARITY).
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BFED7BF443483A CRC64;

Query Match 47.7%; Score 42; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGSRGSRGSGS 19
Db : : : : :
614 SPGSHSQGRGASGS 628

RESULT 9
DXT STRDO STANDARD; PRT; 1337 AA.
ID DXT STRDO STANDARD; PRT; 1337 AA.
AC P39653;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dextranase Precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
DE glucanohydrolase).
GN DEX.
OS Streptococcus downei (Streptococcus sobrinus).
OC Plasmid PYA902.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
RC STRAIN=6715 / UAB66;
RX MEDLINE=94292401; PubMed=8021165;
RA Wanda S.-Y., Curtiss R. III;
RT "Purification and characterization of Streptococcus sobrinus
RT dextranase produced in recombinant Escherichia coli and sequence
RT analysis of the dextranase gene.";
RL J. Bacteriol. 176:3839-3850 (1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
CC PELICLE-COATED TOOTH SURFACE.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
CC linkages in dextran.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
CC and at 39 degrees Celsius.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

DR EMBL; M66378; AAA21772.1; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfam; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KM Hydroxylase; Glycosylase; Cell wall; Peptidoglycan-anchor; Repeat;
KM Signal; Pfamid.
FT CHAIN 1 30
FT PROPEP 1309 1337 DEXTRANASE.
FT SITE 1305 1309 REMOVED BY SORTASE (POTENTIAL).
FT MOD_RES 1308 1308 LPXTG SORTING SIGNAL (POTENTIAL).
SQ SEQUENCE 1337 AA; 14329 MW; B494275A77AE3D0 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 1 SHOEKTXGXRGRSGRS 17
1154 ANQDSTKGSADQSGKS 1170

RESULT 10
WR19 ARATH STANDARD; PRT; 1895 AA.
ID WR19 ARATH
AC Q95Z67;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Probable WRKY transcription factor 19 (WRKY DNA-binding protein 19).
GN WRKY19 OR At4G12020 OR Flj013.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10615198;
RA Meyer K.F.X., Schneller C., Wandut R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Maehre R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hobeisel J., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCallagh B., Bilham L., Robben J.,
RA Van der Schueren U., Grymoprez B., Chuang Y.-J., Vandebussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Wiltzengeser T., Bolte G., Rampsperger U., Hilbert H., Braum M.,
RA Holzer E., Brandt A., Peters S., Van Staveren M., Dirks W.,
RA Moellman P., Klein lankhorst R., Rose M., Hauf J., Koeltter F.,
RA Bernieris S., Hemel S., Feldpausch M., Lambert S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen U., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett U., Hall S., Kay W., Lennard N., McIay K., Mayes R.,
RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderth K., Danner D., Herzl A.,
RA Neumann S., Ayrizrau A., Vitale D., Liguori R., Piravandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Aubourg S.,
RA Schnabl S., Hiller R., Schmidt W., Lechman A., Felter R.,
RA Chefor F., Cooke R., Berger C., Montfort A., Caracuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Punelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bialke C.,
RA Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dehla N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sektion M., Murray J., Saeet P., Cordes M., Abu-Threiden J.,

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Lattrell P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kimer J., Fulton L., Martis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Spieh J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali T., Berhoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Antonov B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekhar M., Madero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman U., Tili S.,
RA Granat S., Shohdy N., Hargreaves A., Hameed A., Iochi M., Johnson A.,
RA Chen E., Marra M.A., Martenssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana";
RL Nature 402:769-777(1999).
CC - FUNCTION: Transcription factor. Interacts specifically with the W
box (5'-(T)GAC(C/T)-3'), a frequently occurring elicitor-
responsive cis-acting element. May act also as a disease
resistance protein with a serine/threonine-protein kinase activity
(By similarity).
CC - SUBCELLULAR LOCATION: Nuclear (Probable).
CC - SIMILARITY: Belongs to the WRKY group I family.
CC - SIMILARITY: Belongs to the disease resistance X-TIR-NB-LRR-X
family.
CC - SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC - SIMILARITY: Contains 1 NB-ARC domain.
CC - SIMILARITY: Contains 1 PAH (paired amphipathic helix) repeat.
CC - SIMILARITY: Contains 1 protein kinase domain.
CC - SIMILARITY: Contains 2 WRKY domains.
CC - DATABASE: NMR-NB-LRRS;
NOTE=Functional and comparative genomics of disease resistance gene
homologs;
WWW="http://nblrire.ucdavis.edu".
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL; AL049638; CAB40943.1; -
CC EMBL; AL161533; CAB78245.1; -
CC PIR; T06609; T06609.
CC HSSP; P02876; 9WGA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR000767; Disease_resist.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR002182; NB-ARC.
DR InterPro: IPR003822; PAH.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR000157; TIR.
DR InterPro: IPR003657; WRKY.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF02671; PAH; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF03106; WRKY; 2.
DR PRINTS; PR00364; DISPARSER1ST.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50811; WRKY; 2.
KM plant defense; transferase; Serine/threonine-protein kinase;
KM transcription regulation; Nuclear protein; ATP-binding; DNA-binding;
KM Repeat; leucine-rich repeat.


```

FT REPEAT 323 369 PAH.
FT DNA_BIND 462 526 WRKY 1.
FT DNA_BIND 635 700 WRKY 2.
FT DOMAIN 800 1087 NB-ARC.
FT REPEAT 1204 1226 LRR 1.
FT REPEAT 1257 1281 LRR 2.
FT REPEAT 1304 1327 LRR 3.
FT REPEAT 1328 1348 LRR 4.
FT REPEAT 1349 1371 LRR 5.
FT REPEAT 1372 1395 LRR 6.
FT REPEAT 1419 1443 LRR 7.
FT DOMAIN 1626 1877 PROTEIN KINASE.
FT NP_BIND 844 851 ATP (POTENTIAL).
FT NP_BIND 1632 1640 ATP (BY SIMILARITY).
FT BINDING 1654 1654 ATP (BY SIMILARITY).
FT ACT_SITE 1758 1758 BY SIMILARITY.
FT DOMAIN 97 307 GLY-RICH.
FT DOMAIN 35 90 POLY-SER.
FT DOMAIN 980 983 POLY-LEU.
FT DOMAIN 1568 1571 POLY-SER.
SQ SEQUENCE 1895 AA; 210320 MW; 1C19D3EE164C9363 CRC64;

```

Query Match 47.2%; Score 41.5; DB 1; Length 1895;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 SHOESTYXGRSGS 15
 Db 219 SHEDCT-GPARCRSG 232

```

RESULT 11
SF5_RAT 585_RAT STANDARD; PRT; 825 AA.
AC 063003;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE 5E5 antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein 5E5 in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, D37934; BAA07153.1; -.
DR PIR; JC4163; JC4163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 8681 MW; AF667FE2F555BDF CRC64;

```

Query Match 46.6%; Score 41; DB 1; Length 825;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GXRGGRSGS 19
 Db 589 GGRGRGGRSGS 600

```

RESULT 12
PRO_CRESP ID PRO_CRESP STANDARD; PRT; 1015 AA.
AC Q00586;
DT 15-JUL-1999 (Rel. 38; Created)
DT 15-JUL-1999 (Rel. 38; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Frequency clock protein.
DE PRO.
OS Creopus spinulosus (Chromocera spinulosa).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=110619;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4896;
RX MEDLINE=97188515; PubMed=9037100;
RA Lewis M.T., Morgan L.W., Feldman J.F.;
RT "Analysis of frequency (frq) clock gene homologs: evidence for a
RT helix-turn-helix transcription factor.";
RL Mol. Gen. Genet. 253:401-414(1997).
CC -1- FUNCTION: Circadian clock component involved in the generation of
CC biological rhythms, in particular in rhythm stability period
CC length, and temperature compensation. Behaves as a negative
CC element in circadian transcriptional loop (by similarity).
CC -1- SIMILARITY: BELONGS TO THE FRQ FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U25850; AAA6072.1; -.
DR PIR; T42013; T42013.
KW Biological rhythms; Transcription regulation; Nuclear protein.
FT DOMAIN 240 245 POLY-SER.
FT DOMAIN 356 368 POLY-SER.
FT DOMAIN 443 451 POLY-SER.
FT DOMAIN 584 588 POLY-SER.
FT DOMAIN 892 913 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 1015 AA; 110972 MW; EA9E732ED7414B1 CRC64;

```

Query Match 46.6%; Score 41; DB 1; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 43;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTYXGRSGS 19
 Db 583 NHRQKGTGSTGSSGN 601

```

RESULT 13
COAC_YEAST ID COAC_YEAST STANDARD; PRT; 2233 AA.
AC Q00955;
DT 01-APR-1993 (Rel. 25; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.3.4.14).
DE DE CARBOXYLASE (EC 6.3.4.14).
DE PAS3 OR ACC1 OR YNR015C OR N3175.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

```

RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2015-2022.
RX MEDLINE=92262474; PubMed=1350093;
RA Al-Feel W., Chitrals S.S., Wakil S.J.;
RT "Cloning of the yeast P433 gene and primary structure of yeast
RL Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Submitted (MAY.1996) to the EMBL/GenBank/DBJ databases.
RL Pohl T.M.;
CC -1- FUNCTION: This protein carries three functions: biotin carboxyl
CC carrier protein, biotin carboxylase, and carboxyltransferase.
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: Biotin.
CC -1- ENZYME REGULATION: By phosphorylation.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EXBL: M92156; AAA20073.1; -
DR EMBL: Z71631; CA96294.1; -
DR PIR: S63347; S63347.
DR HSSP: P24182; IBNC.
DR Germonline; 143361; -
DR SGD: S0005299; ACCL.
DR GO: GO:0005789; Cytoplasmic reticulum membrane; IDA.
DR GO: GO:0003989; Fatty acid biosynthesis; IMP.
DR GO: GO:0004075; Fatty acid biosynthesis; IMP.
DR InterPro: IPR001882; Biotin_BS.
DR InterPro: IPR005482; Biotin_carb_C.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000022; Carboxyl_trans.
DR InterPro: IPR005479; CPase_L_D2.
DR InterPro: IPR005481; CPase_L_N.
DR Pfam: PF02785; Biotin_carb_C_1.
DR Pfam: PF00364; biotin_lipoyl_1.
DR Pfam: PF01039; Carboxyl_trans_1.
DR Pfam: PF02789; CPase_L_trans_1.
DR Pfam: PF02786; CPase_L_D2_1.
DR PROSITE: PS00188; BIOTIN; 1.
DR PROSITE: PS00866; CPASE_1; 1.
DR PROSITE: PS00867; CPASE_2; 1.
KM Fatty acid biosynthesis; Biotin ligase; Multifunctional enzyme;
KM ATP-binding; Phosphorylation.
FT NP_BIND 256 261 ATP (BY SIMILARITY).
FT ACT_SITE 383 383 BY SIMILARITY.
FT BINDING 735 735 BIOTIN (BY SIMILARITY).
FT DOMAIN 1865 1894 COENZYME A-BINDING (BY SIMILARITY).
FT CONFLICT 1523 1523 W -> G (IN REF. 1).
FT CONFLICT 1755 1755 I -> TWYGL (IN REF. 1).
FT CONFLICT 1761 1766 AINRM -> ESTNA (IN REF. 1).
SQ SEQUENCE 2233 AA; 250351 MW; 0A335A09B1F8308 CRC64;
Query Match 56.6%; Score 41; DB 1; Length 2233;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 2 HQSTGXSRGSRGSGS 19

DB 1200 HQSSNCPADPRGSSGS 1217
RESULT 14
ID ROA3 MOUSE STANDARD; PRT; 379 AA.
AC Q8BG05; Q8BHF8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3).
GN HNRP3 OR HNRP3.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Baysaralain D.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=FVB/N; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Dietchenko L., Marinska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stedman M., Soares M.B., Bonaldo M.P., Casavant J.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Pange C.,
RA Raskas S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8BG05-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8BG05-2; Sequence=VSP_007350.
CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRP, WHICH ARE BASIC AND
CC GUY-RICH PROTEINS.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RNM) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF463524; AAN76922.1; -
DR EMBL: BC023828; AAH23828.1; -
DR EMBL: BC023908; AAH23908.1; -
DR EMBL: BC038364; AAH38364.1; -
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm_2.
DR SMART: SM00360; RRM_2.
DR PROSITE: PS50102; RRM_2.
DR PROSITE: PS00030; RRM_RNP_1; 2.

KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
 KM Alternative splicing. 35 118 RNA-BINDING (RRM) 1.
 FT DOMAIN 126 205 RNA-BINDING (RRM) 2.
 FT DOMAIN 211 379 GLY-RICH.
 FT VARSPLIC 1 23 MEVRRPPGRPQDSGRRRRRRGRG -> M (in isoform 2).
 FT
 FT SEQUENCE 379 AA; 39652 MW; D83C400A2B096E95 CRC64;
 Query Match Best Local Similarity 46.0%; Score 40.5; DB 1; Length 379;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;
 QY 1 SHQSTGXSRGRSGSGS 19
 Db 203 SKQKQSAQSQRGRGSGSGN 222
 RESULT 15
 DH1_MAIZE STANDARD; PRT; 168 AA.
 AC P12550; P16318; (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Dehydrin DHN1 (M3) (RAB-17 protein).
 GN DHN1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93357436; PubMed=2562763;
 RA Close F.J., Korte A.A., Chandler P.M.;
 RT "A cDNA-based comparison of dehydration-induced proteins (dehydrins)
 in barley and corn."
 RL Plant Mol. Biol. 13:95-108(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV.AC.1503;
 RX MEDLINE=91346629; PubMed=2151715;
 RA Vilardell J., Goday A., Freire M.A., Torrent M., Martinez M.C.,
 RA Torne J.M., Pages M.;
 RT "Gene sequence, developmental expression, and protein phosphorylation
 of RAB-17 in maize."
 RL Plant Mol. Biol. 14:423-432(1990).
 RN [3]
 RP PHOSPHORYLATION BY CKII.
 RX MEDLINE=92042196; PubMed=1939268;
 RA Plana M., Itarte E., Ertija R., Goday A., Pages M., Martinez M.C.;
 RT "Phosphorylation of maize RAB-17 protein by casein kinase 2."
 RL J. Biol. Chem. 266:22510-22514(1991).
 CC -1- DEVELOPMENTAL STAGE: ABA treatment induced the synthesis of RAB-17
 in calli, however, the RAB-17 proteins were found to be highly
 phosphorylated only in embryos.
 CC -1- INDUCTION: By abscisic acid (ABA) and water stress.
 CC -1- PTM: SERINE IS THE ONLY AMINO ACID WHICH BECOMES PHOSPHORYLATED.
 CC -1- SIMILARITY: Belongs to the plant dehydrin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X15290; CAA33364.1; -;
 DR EMBL: X15994; CAA34123.1; -;
 DR PIR: A39316; A39316.

DR PIR: S05545; S05545.
 DR PIR: S08633; S08633.
 DR MaizeDB; 24916; -;
 DR InterPro; IPR000167; Dehydrin.
 DR Pfam; PF00257; dehydrin; 1.
 DR PROSITE; PS00315; DEHYDRIN 1; 1.
 DR PROSITE; PS00823; DEHYDRIN 2; 2.
 KW Dehydrin; Repeat; Phosphorylation; Multigene family.
 FT DOMAIN 76 84 POLY-SER.
 FT DOMAIN 2 140 3 X APPROXIMATE REPEATS.
 FT REPEAT 2 10 1-1.
 FT REPEAT 122 130 1-2.
 FT REPEAT 132 140 1-3.
 FT DOMAIN 94 166 2 X APPROXIMATE REPEATS.
 FT REPEAT 94 107 2-1.
 FT REPEAT 153 166 2-2.
 FT CONFLICT 9 9 R -> H (IN REF. 2).
 FT CONFLICT 13 13 R -> A (IN REF. 2).
 FT CONFLICT 41 41 MISSING (IN REF. 2).
 SQ SEQUENCE 168 AA; 17161 MW; BE7758CAD37CFE39 CRC64;
 Query Match Best Local Similarity 45.5%; Score 40; DB 1; Length 168;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 2 HQSTGXSRGRSGSGS 19
 Db 114 HATATTGAYGQGGHTGS 131

Search completed: September 28, 2004, 06:05:50
 Job time : 8.22396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-6
Sequence: 1 SH0ESTYXGSRGSRGSRGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	797	4	Q16824
2	84	95.5	990	4	Q15206
3	84	95.5	1218	4	Q05331
4	78	88.6	798	4	Q9H4U3
5	78	88.6	1084	4	Q01212
6	69	78.4	591	4	Q01720
7	69	78.4	687	4	Q9H4U2
8	67	76.1	322	4	Q75370
9	64	72.7	465	4	Q03838
10	49	55.7	822	3	Q9P312
11	47	53.4	96	15	Q8JBU5
12	47	53.4	96	15	Q9WRF4
13	47	53.4	96	15	Q8JBU4
14	47	53.4	96	15	Q8ADFS
15	47	53.4	399	4	Q9NTA9
16	47	53.4	467	4	Q9BWT8

17	47	53.4	700	13	Q42378	Q42378 brachydanio
18	47	53.4	715	13	Q8QF00	Q8QF00 brachydanio
19	47	53.4	715	13	Q8QGG8	Q8QGG8 brachydanio
20	47	53.4	716	13	Q42107	Q42107 brachydanio
21	47	53.4	737	4	Q9BQ39	Q9BQ39 homo sapien
22	47	53.4	843	12	Q9QNS2	Q9QNS2 hepatitis b
23	46	52.3	96	15	Q8ADX3	Q8ADX3 human immun
24	46	52.3	243	10	Q9LX45	Q9LX45 oryza sativ
25	46	52.3	426	4	Q8W776	Q8W776 homo sapien
26	45	51.1	669	11	Q8BRU5	Q8BRU5 mus musculu
27	45	51.1	734	11	Q99M19	Q99M19 mus musculu
28	44.5	50.6	486	10	Q7XQ91	Q7XQ91 oryza sativ
29	44	50.0	96	15	Q9WSQ0	Q9WSQ0 human immun
30	44	50.0	96	15	Q9W9K9	Q9W9K9 human immun
31	44	50.0	96	15	Q90CLO	Q90CLO human immun
32	44	50.0	96	15	Q9WSP4	Q9WSP4 human immun
33	44	50.0	96	15	Q998H5	Q998H5 human immun
34	44	50.0	96	15	Q72615	Q72615 human immun
35	44	50.0	96	15	Q90CK1	Q90CK1 human immun
36	44	50.0	96	15	Q9WRG0	Q9WRG0 human immun
37	44	50.0	96	15	Q9WRE7	Q9WRE7 human immun
38	44	50.0	96	15	Q8UMH2	Q8UMH2 human immun
39	44	50.0	96	15	Q91MJ5	Q91MJ5 human immun
40	44	50.0	96	15	Q9Q6H4	Q9Q6H4 human immun
41	44	50.0	96	15	Q72610	Q72610 human immun
42	44	50.0	96	15	Q90DY9	Q90DY9 human immun
43	44	50.0	96	15	Q9QGX5	Q9QGX5 human immun
44	44	50.0	96	15	Q90DZ9	Q90DZ9 human immun
45	44	50.0	103	12	Q9WMD4	Q9WMD4 hepatitis b

ALIGNMENTS

RESULT 1
Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9106347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene [published erratum appears in Biochemistry 1991 Jun
11;30(23):5814].";
RT 11,30(23):5814. ;
RL Biochemistry 29:9432-9440(1990).
DR EMBL: M60502; AAA63248.1; -
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro:IPR003303; F:laggrin.
DR Pfam: PF0516; F:laggrin; 4.
FT PRINTS; PR00487; F:LAGGRIN.
FT NON_TER 1
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDA86B CRC64;
Query Match 95.5%; Score 84; DB 4; Length 797;
Best Local Similarity 89.5%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SH0ESTYXGSRGSRGSRGS 19
|||
DB 427 SH0ESTYXGSRGSRGSRGS 445

RESULT 2

```

015206
ID 015206 PRELIMINARY; PRT; 990 AA.
AC 015206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
FT NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGSGRSGSGS 19
Db 551 SHOESTRGSGRSGSGS 569

RESULT 3
ID 005331 PRELIMINARY; PRT; 1218 AA.
AC 005331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993)
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA6487.1; -.
DR PIR; A48118; A48118.
DR HSSP; P02593; ICDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CAAP; 1.
KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGSGRSGSGS 19
Db 773 SHOESTRGSGRSGSGS 791

RESULT 4
ID 094U3 PRELIMINARY; PRT; 798 AA.
AC 094U3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DJ4NL.1.2 (Profilaggrin 3' end) (Fragment).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
[1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 88.6%; Score 78; DB 4; Length 798;
Best Local Similarity 84.2%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGSGRSGSGS 19
Db 428 SHOESTRGSGRSGSGS 446

RESULT 5
ID 001212 PRELIMINARY; PRT; 1084 AA.

```

```

AC Q01212; Q03840;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814]";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON_TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640BDEA362D CRC64;

Query Match 88.6%; Score 78; DB 4; Length 1084;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOE2TXGSRGSRSGS 19
DB 64 SHOE2TXGSRGSRSGS 82

RESULT 6
Q01720 PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Filaggrin precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Nirsunskaiti W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus.";
RL J. Biol. Chem. 267:23772-23781(1992).
U.
-i- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFIDE-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
-i- PTM: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
PROTEOLYTICALLY CLEAVED.
-i- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135840; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293 POTENTIAL.
FT CHAIN 294 467 FILAGRIN.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 POTENTIAL.
FT CA BIND 19 32 FILAGRIN.
FT CA BIND 62 73 SITE I (BY SIMILARITY).
FT NON_TER 591 591 SITE II (BY SIMILARITY).
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75B369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.003;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOE2TXGSRGSRSGS 19
DB 449 SHOE2TXGSRGSRSGS 467

RESULT 7
Q094U2 PRELIMINARY; PRT; 687 AA.
AC Q094U2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DJ4N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
U.
-i- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.0035;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 SHQESTXGXSRRGSRGSGS 19
      |||||
Db      449 SHQESTRGRGSRGSGSGS 467

RESULT 8
ID 075370 PRELIMINARY; PRT; 322 AA.
AC 075370;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Epidermal filaggrin (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9101527; PubMed=9886436;
RA Girbal-Neuhausser E., Durieux J.J., Arnaud M., Dalbon P., Sebba M.,
RA Vincent C., Simon M., Senhu T., Masson-Bessiere C.,
RA Jolivet-Reynaud C., Jolivet M., Serre G.;
RT "The epitopes targeted by the rheumatoid arthritis-associated
RT anti-filaggrin autoantibodies are posttranslationally generated on
RT various sites of (pro)filaggrin by deamination of arginine residues.";
RL J. Immunol. 162:585-584(1999).
DR EMBL; AF043380; AAC23559.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 2.
DR PRINTS; PRO0487; FILAGRIN.
FT NON_TER 1
FT NON_TER 322
SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0330DBF92E0 CRC64;

Query Match
Best Local Similarity 76.1%; Score 67; DB 4; Length 322;
Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 SHQESTXGXSRRGSRGSG 18
      |||||
Db      305 SHQESTRGRGSRGSGSG 322

RESULT 9
ID 003838 PRELIMINARY; PRT; 465 AA.
AC 003838;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGRIN (PROFILAGRIN) (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP REVISIONS
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814(1991).

```

```

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
CC REPEATS.
CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE
CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalIN
CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
DR EMBL; M60499; AAA63246.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 3.
DR PRINTS; PRO0487; FILAGRIN.
FT NON_TER 1
FT NON_TER 465
SQ SEQUENCE 465 AA; 50280 MW; C883744C5E1334097 CRC64;

Query Match
Best Local Similarity 72.7%; Score 64; DB 4; Length 465;
Pred. No. 0.016;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 SHQESTXGXSRRGSRGSGS 19
      |||||
Db      227 SHQESARGRGRGSRGSGS 245

RESULT 10
ID 09P312 PRELIMINARY; PRT; 822 AA.
AC 09P312;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12F1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aligh V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR005120; Smg-4_UZF3.
DR Pfam; PF00076; trm; 1.
DR Pfam; PF03467; Smg4_UZF3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match
Best Local Similarity 55.7%; Score 49; DB 3; Length 822;
Pred. No. 9.8;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      3 QESTXGXSRRGSRGSGS 19
      |||||
Db      414 RESASGRTRGRGRGCT 430

RESULT 11

```


Q8JBUS
ID Q8JBUS PRELIMINARY; PRT; 96 AA.
AC Q8JBUS;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=00KE_XH1199;
RA Dowling W.E., Kim B., Mason C.J., Masuna K.Monique., Alam U.,
RA Elson L., Bix D.L., Robb M.L., McCutchan F.E., Carr J.K.;
RT "Forty-one near full length HIV-1 sequences from Kenya reveal an
epidemic of subtype A and A-containing recombinants."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457067; AAN03148.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11300 MW; 4A60A702801D3473 CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

RESULT 12
Q9WRP4
ID Q9WRP4 PRELIMINARY; PRT; 96 AA.
AC Q9WRP4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96CWR1000.4/96CWR1.4;
RX MEDLINE=99329206; PubMed=10400779;
RA Takenaka U., Zekeng L., Ido E., Yamaguchi-Kabata Y., Mboudjeka I.,
RA Harada Y., Miura T., Kapu L., Hayami M.;
RT "Human immunodeficiency virus type 1 intergroup (M/O) recombination in
cancer.";
RL J. Virol. 73:6810-6820(1999).
DR EMBL; AF097693; AAD41705.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11254 MW; F01C751229CFB9A CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 52.9%; Pred. No. 2.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGA 93

RESULT 13

Q8JBV4
ID Q8JBV4 PRELIMINARY; PRT; 96 AA.
AC Q8JBV4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=00KE_XH1144;
RA Dowling W.E., Kim B., Mason C.J., Masuna K.Monique., Alam U.,
RA Elson L., Bix D.L., Robb M.L., McCutchan F.E., Carr J.K.;
RT "Forty-one near full length HIV-1 sequences from Kenya reveal an
epidemic of subtype A and A-containing recombinants."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457066; AAN03139.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11319 MW; A23861A05083196B CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

RESULT 14
Q8ADP5
ID Q8ADP5 PRELIMINARY; PRT; 96 AA.
AC Q8ADP5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98UG57142;
RX MEDLINE=22375625; PubMed=12487816;
RA Harris K.E., Serwadda D., Sewankambo N., Wabwire F., Kim B.,
RA Kigozi G., Kiwanuka N., Phillips J.B., Mehen M., Lutalo T.,
RA Laey J.R., Merling R., Gray R., Waver M., Bix D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 Near Full Length HIV Type 1 Genome Sequences from Rakai
District, Uganda, Subtype D and AD Recombinants Predominate."
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
DR EMBL; AF484512; AAN73753.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11356 MW; 5EAB1F03A30F5221 CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

```

RESULT 15
O9NTA9
ID O9NTA9          PRELIMINARY;      PRT;      399 AA.
AC O9NTA9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZ76130333.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Bioecker H., Bioecker M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.,
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137423; CAB70733.1; -.
DR PIR; T46259; T46259.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:Nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00490; HEHIC; 1.
KM Hypothetical protein; ATP-binding; Helicase; Hydrolase.
FT NON TER
SQ SEQUENCE 399 AA; 44075 MW; BC09B3A0E4CE5E96 CRC64;

Query Match      53.4%; Score 47; DB 4; Length 399;
Best Local Similarity 55.6%; Pred. No. 9.9;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Cy 1 SHQSTXGKSRGRSGRSG 18
   | | | | | | | | | |
Db 336 SNGRSGWSSGRSGRSG 353

```

Search completed: September 28, 2004, 06:12:45
 Job time : 41.1771 secs

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 6e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGXSRRGRSGSGS 19
 |||||
 1 SHOESTAGRSRGRSGSGS 19

RESULT 2

ID AAW61506 standard; peptide; 19 AA.

XX AAW61506;

DT 26-OCT-1998 (first entry)

DE Peptide cfc2, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note="Citrulline"

PN W09822503-A2.

XX 28-MAY-1998.

PF 14-NOV-1997; 97WO-NI000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKINDIG ONDERZOEK IN NEDER.

PI (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

SQ

QY

DB

1 SHOESTXGXSRRGRSGSGS 19

1 SHOESTAGRSRGRSGSGS 19

AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

DT 26-OCT-1998 (first entry)

DE Peptide cfc2, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

PN W09822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NI000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKINDIG ONDERZOEK IN NEDER.

PI (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

SQ

QY

DB

1 SHOESTXGXSRRGRSGSGS 19

1 SHOESTAGRSRGRSGSGS 19

RESULT 4

AAW61510

ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

DT 26-OCT-1998 (first entry)

DE Peptide cfc6, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 9 /note= "Citruilline"
FT Modified-site 9 /note= "Citruilline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SHOESTXGSRGSRGSRGS 19
DB 1 SHOESTXGSRGSRGSRGS 19
XX
RESULT 5
AAM61505
ID AAM61505 standard; peptide; 19 AA.
XX
AC AAM61505;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cft, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
OS Synthetic.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 7 /note= "Citruilline"
XX
XX WO9822503-A2.
XX 28-MAY-1998.
XX

```

```

XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.7e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHOESTXGSRGSRGSRGS 19
DB 1 SHOESTXGSRGSRGSRGS 19
XX
RESULT 6
AAM61514
ID AAM61514 standard; peptide; 19 AA.
XX
AC AAM61514;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cf, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
OS Synthetic.
XX Homo sapiens.
XX
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX

```

PS Disclosure; Page 6; 19pp; English.

XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX Sequence 19 AA;

XX

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.7e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQESTYXGSRGRSGSGS 19

Db 1 SHQESTYXGSRGRSGSGS 19

RESULT 7

AAW61516

ID AAM61516 standard; peptide; 19 AA.

XX

XX AAM61516;

XX

XX 26-OCT-1998 (first entry)

XX

DE Peptide cFE, based on cDNA of a proflilagrin repeat.

XX

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX

XX monoclonal antibody.

XX

XX Synthetic.

XX

OS Homo sapiens.

XX

XX W09822503-A2.

XX

XX 28-MAY-1998.

XX

XX

XX 14-NOV-1997; 97MO-NL000624.

XX

XX

XX 15-NOV-1996; 96NL-01004539.

XX

XX (SCHE-) STICHTING SCHEIKONDIC ONDERZOEK IN NEDER.

XX

XX (TEWE-) STICHTING TECH WETENSCHAPEN.

XX

XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;

XX

XX WPI; 1998-398613/34.

XX

XX

XX Peptide derived from an antigen recognised by autoantibodies - is

XX

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX

XX used in diagnosis of the disease.

XX

XX

PS Disclosure; Page 6; 19pp; English.

XX

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of

XX

XX the proflilagrin antigen which is recognised by autoantibodies from

XX

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX

XX RA patient's autoimmune antibodies which are reactive with proflilagrin.

XX

XX The peptides were created by using standard solid phase synthesis, which

XX

XX produced them as peptide amides. These sequences may be used in the

XX

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX

XX for obtaining polyclonal and monoclonal antibodies

XX

XX Sequence 19 AA;

XX

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.7e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQESTYXGSRGRSGSGS 19

Db 1 SHQESTYXGSRGRSGSGS 19

RESULT 8

ABB97605

ID ABB97605 standard; protein; 1467 AA.

XX

XX ABB97605;

XX

XX 27-JUN-2002 (first entry)

XX

XX

XX Novel human protein SEQ ID NO: 873.

XX

XX

XX Human; antihaemic; vulnery; antiinflammatory; immunomodulator;

XX

XX antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;

XX

XX neuroprotective; antiparkinsonian; protein therapy; EST;

XX

XX expressed sequence tag.

XX

XX Homo sapiens.

XX

XX W0200222660-A2.

XX

XX

XX 21-MAR-2002.

XX

XX

XX 10-SEP-2001; 2001MO-US026015.

XX

XX

XX 11-SEP-2000; 2000US-00659671.

XX

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX

XX Xue AJ, Yang Y, Wehrman T, Dymnac RT;

XX

XX WPI; 2002-292408/33.

XX

XX N-PSDB; ABN32791.

XX

XX

XX An isolated polynucleotide for treating diseases associated with its

XX

XX encoded polypeptide such as cancer and multiple sclerosis.

XX

XX

XX Example 2; SEQ ID NO 873; 509pp; English.

XX

XX

XX The present invention provides the protein and coding sequences of 444

XX

XX novel human proteins. These were isolated from expressed sequences tags

XX

XX (ESTs). They can be used to stimulate cell growth, to regulate

XX

XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

XX

XX e.g. in burn treatment, to regulate the immune system e.g. to treat

XX

XX multiple sclerosis, to regulate activin or inhibin e.g. to treat

XX

XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

XX

XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.

XX

XX rheumatoid arthritis, and to treat nervous system disorders e.g.

XX

XX Parkinson's disease. The present sequence is a protein of the invention

XX

XX

XX Sequence 1467 AA;

XX

Query Match 95.5%; Score 84; DB 5; Length 1467;

Best Local Similarity 89.5%; Pred. No. 6e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQESTYXGSRGRSGSGS 19

Db 773 SHQESTYXGSRGRSGSGS 791

RESULT 9

AAW61508

ID AAM61508 standard; peptide; 19 AA.

XX

XX AAM61508;

XX

DT 26-OCT-1998 (first entry)
 XX Peptide cfc4, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 8.2e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHOESTXGXRGRSGSGS 19
 DB 1 SHOESTRGRXRGRSGSGS 19
 RESULT 10
 AAM61507
 ID AAM61507 standard; peptide; 19 AA.
 AC AAM61507;
 XX
 XX 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc3, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 11
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 8.2e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHOESTXGXRGRSGSGS 19
 DB 1 SHOESTRGRXRGRSGSGS 19
 RESULT 11
 AAM61512
 ID AAM61512 standard; peptide; 19 AA.
 AC AAM61512;
 XX
 XX 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruiline"
 FT Modified-site 13
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX

```

XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 8.2e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 SHOESTXGXRGRSGRSGS 19
XX 1 SHOESTXGXRGRSGRSGS 19
XX
XX RESULT 12
XX AAM61509
XX ID AAM61509 standard; peptide; 19 AA.
XX
XX AAM61509;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc5, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 16 /note= "Citruilline"
XX FT
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX

```

```

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 84.2%; Pred. No. 8.2e-06;
XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 SHOESTXGXRGRSGRSGS 19
XX 1 SHOESTXGXRGRSGRSGS 19
XX
XX Db
XX
XX RESULT 13
XX AAM61511
XX ID AAM61511 standard; peptide; 19 AA.
XX
XX AAM61511;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 7 /note= "Citruilline"
XX FT Modified-site 11 /note= "Citruilline"
XX FT
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX

```


CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 8.2e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGSRGRSGSGS 19
 |||||
 DB 1 SHOESTXGSRGRSGSGS 19

RESULT 14

AAW61513
 ID AAW61513 standard; peptide; 19 AA.

XX AC AAW61513;

XX DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonaal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 16 /note= "Citruilline"

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MWM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 8.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SHOESTXGSRGRSGSGS 19
 |||||
 DB 1 SHOESTXGSRGRSGSGS 19

RESULT 15

AAW61520
 ID AAW61520 standard; peptide; 21 AA.

XX AC AAW61520;

XX DT 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonaal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MWM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 2.2e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HOESTXGSRGRSGSGS 19
 |||||
 DB 4 HOESTXGSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:16
 Job time : 55.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-6

Perfect score: 88

Sequence: 1 SHQESTYXGSRGRSGRSGS 19

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-2
4	84	95.5	19	9	US-09-308-150-6
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	78	88.6	19	9	US-09-308-150-3
9	78	88.6	19	9	US-09-308-150-4
10	78	88.6	19	9	US-09-308-150-5
11	78	88.6	19	9	US-09-308-150-7
12	78	88.6	19	9	US-09-308-150-8
13	78	88.6	19	9	US-09-308-150-9
14	75	85.2	21	9	US-09-308-150-10
15	75	85.2	22	9	US-09-747-029A-22

16	49	55.7	506	16	US-10-437-963-185974	Sequence 185974,
17	46	52.3	243	16	US-10-437-963-103033	Sequence 103033,
18	44.5	50.6	181	16	US-10-437-963-117640	Sequence 117640,
19	44	50.0	155	12	US-10-425-114-40476	Sequence 40476, A
20	44	50.0	435	12	US-10-425-114-62866	Sequence 62866, A
21	44	50.0	430	16	US-10-437-963-164039	Sequence 164039,
22	44	50.0	477	15	US-10-161-927-62	Sequence 62,
23	44	50.0	653	9	US-09-746-801A-13	Sequence 13,
24	44	50.0	730	14	US-10-359-431-19	Sequence 49,
25	44	50.0	842	14	US-10-359-431-50	Sequence 51,
26	44	50.0	842	14	US-10-359-431-51	Sequence 50,
27	44	50.0	843	14	US-10-209-264-2	Sequence 2,
28	44	50.0	843	14	US-10-359-431-45	Sequence 45,
29	44	50.0	843	14	US-10-359-431-48	Sequence 48,
30	44	50.0	843	14	US-10-359-431-49	Sequence 49,
31	44	50.0	845	14	US-10-359-431-56	Sequence 56,
32	44	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
33	44	50.0	1087	9	US-09-918-909-24	Sequence 24,
34	44	50.0	1087	16	US-10-641-991-74	Sequence 74,
35	43	48.9	232	16	US-10-437-963-183394	Sequence 183394,
36	43	48.9	247	12	US-10-424-589-187446	Sequence 187446,
37	43	48.9	455	16	US-10-437-963-185897	Sequence 185897,
38	43	48.9	484	16	US-10-437-963-185432	Sequence 185432,
39	43	48.9	488	16	US-10-437-963-146097	Sequence 146097,
40	43	48.9	570	10	US-09-847-102A-43	Sequence 43,
41	43	48.9	694	14	US-10-152-548-2	Sequence 2,
42	43	48.9	797	14	US-10-032-585-7208	Sequence 7208, Ap
43	42.5	48.3	92	16	US-10-437-963-118183	Sequence 118183,
44	42	47.7	34	9	US-09-864-761-37121	Sequence 37121, A
45	42	47.7	179	12	US-10-425-114-53545	Sequence 53545, A

ALIGNMENTS

US-09-308-150-12
; Sequence 12, Application US/09308:150
; Patent No. US20020137092A1
GENERAL INFORMATION:
; APPLICANT: Van Veenrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: US/09/308,150
PRIOR APPLICATION NUMBER: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-09-308-150-12
OTHER INFORMATION: Known CDNA sequences of human profilaggrin
Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.3e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 SHOESTXGSRGSRGSRGS 19

RESULT 2

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308,150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 7.5e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGSRGSRGSRGS 19
 Db 1 SHOESTXGSRGSRGSRGS 19

RESULT 3

US-09-308-150-2
 ; Sequence 2, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308,150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 19

TYPE: PRT
 ORGANISM: Artificial Sequence

FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-2

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 7.5e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGSRGSRGSRGS 19
 Db 1 SHOESTXGSRGSRGSRGS 19

RESULT 4

US-09-308-150-6
 ; Sequence 6, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308,150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-6

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 100.0%; Pred. No. 7.5e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHOESTXGSRGSRGSRGS 19
 Db 1 SHOESTXGSRGSRGSRGS 19

RESULT 5

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-11
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGXSRRGRSGSGS 19
Db      1 SHOESTGRGRSGRSGSGS 19
```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-13
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGXSRRGRSGSGS 19
Db      1 SHOESTGRGRSGRSGSGS 19
```

```
RESULT 7
US-09-308-150-14
```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-14
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGXSRRGRSGSGS 19
Db      1 SHOESTGRGRSGRSGSGS 19
```

```

RESULT 8
US-09-308-150-3
; Sequence 3, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3
```

Query Match 88.6%; Score 78; DB 9; Length 19;
 Best Local Similarity 84.2%; Pred. No. 5.9e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSGSGS 19
 Db 1 SH0ESTRGRSGRSGSGS 19

RESULT 9

US-09-308-150-4
 Sequence 4, Application US/09308150
 Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 APPLICANT: Stichting voor de Technische Wetenschappen
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: NL 1004539
 PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-4

Query Match 88.6%; Score 78; DB 9; Length 19;
 Best Local Similarity 84.2%; Pred. No. 5.9e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSGSGS 19
 Db 1 SH0ESTRGRSGRSGSGS 19

RESULT 10

US-09-308-150-5
 Sequence 5, Application US/09308150
 Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 APPLICANT: Stichting voor de Technische Wetenschappen
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624
 PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
 PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 5
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
 Best Local Similarity 84.2%; Pred. No. 5.9e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSGSGS 19
 Db 1 SH0ESTRGRSGRSGSGS 19

RESULT 11

US-09-308-150-7
 Sequence 7, Application US/09308150
 Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 APPLICANT: Stichting voor de Technische Wetenschappen
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: NL 1004539
 PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 7
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-7

Query Match 88.6%; Score 78; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.9e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSGSGS 19
 Db 1 SH0ESTXGXSRRGSGSGS 19

RESULT 12

US-09-308-150-8
 Sequence 8, Application US/09308150
 Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-8

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGSRGRSGSGS 19
Db 1 SHQSTXGSRGRSGSGS 19

RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308:150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-9

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGSRGRSGSGS 19
Db 1 SHQSTXGSRGRSGSGS 19

RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308:150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
NAME/KEY: DISULFID
LOCATION: (3) .. (16)
US-09-308-150-10

Query Match 85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.00018;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGSRGRSGSGS 19
Db 4 HOESTXGSRGRSGSGS 21

RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747:029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Arn
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydia
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.0031NPOS00 INNS:031
CURRENT APPLICATION NUMBER: US/09/747,029A
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptides
NAME/KEY: MOD_RES
LOCATION: (10)..(10)
OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

Query Match 85.2%; Score 75; DB 9; Length 22;
Best Local Similarity 88.9%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGXSRRGRCGRSGS 19
|||
Db 5 HOESTXGXSRRGRCGRSGS 22

Search completed: September 28, 2004, 07:28:51
Job time : 112.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-6

Perfect score: 88

Sequence: 1 SHQSTXKXSRGSRGSGS 19

Scoring table: BLOSUM62

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfill.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	54.5	421	4	US-09-252-991A-32326
2	44	50.0	483	4	US-09-252-991A-19015
3	44	50.0	653	4	US-09-513-057C-13
4	44	50.0	730	4	US-08-591-502B-49
5	44	50.0	842	4	US-08-591-502B-50
6	44	50.0	842	4	US-08-591-502B-51
7	44	50.0	843	4	US-09-719-528A-2
8	44	50.0	843	4	US-08-591-502B-45
9	44	50.0	843	4	US-08-591-502B-48
10	44	50.0	843	4	US-08-591-502B-59
11	44	50.0	845	4	US-08-591-502B-46
12	43	48.9	169	3	US-09-342-084-6
13	42	47.7	138	4	US-09-252-991A-16961
14	42	47.7	782	4	US-09-252-991A-24514
15	42	47.7	782	4	US-09-252-991A-31219
16	42	47.7	724	4	US-09-884-890-2
17	42	47.7	745	4	US-09-523-849-36
18	41	46.6	255	4	US-09-252-991A-19598
19	41	46.6	258	4	US-09-252-991A-19857
20	41	46.6	406	4	US-09-252-991A-19857
21	41	46.6	409	4	US-09-489-039A-14217
22	41	46.6	464	4	US-09-252-991A-31219
23	41	46.6	722	4	US-08-817-832B-32
24	41	46.6	2237	4	US-08-354-973-1
25	40.5	46.0	258	4	US-10-164-595-59
26	40.5	46.0	274	4	US-09-976-594-417
27	40.5	46.0	378	4	US-10-164-595-2

28	40	45.5	120	4	US-09-252-991A-17025	Sequence 17025, A
29	40	45.5	197	4	US-09-252-991A-23579	Sequence 23579, A
30	40	45.5	274	4	US-09-252-991A-16696	Sequence 16696, A
31	40	45.5	465	4	US-09-252-991A-16919	Sequence 16919, A
32	40	45.5	486	2	US-08-821-355A-8	Sequence 8, Appl1
33	40	45.5	486	2	US-09-003-687A-8	Sequence 8, Appl1
34	40	45.5	486	2	US-09-136-605-8	Sequence 8, Appl1
35	40	45.5	508	2	US-08-818-024-3	Sequence 3, Appl1
36	40	45.5	508	3	US-09-334-725A-3	Sequence 3, Appl1
37	40	45.5	660	3	US-09-058-489-18	Sequence 18, Appl1
38	40	45.5	660	3	US-09-058-489-91	Sequence 91, Appl1
39	40	45.5	660	4	US-09-976-594-787	Sequence 787, Appl1
40	40	45.5	708	4	US-09-252-991A-16074	Sequence 16074, A
41	39.5	44.9	160	4	US-09-252-991A-26901	Sequence 26901, A
42	39	44.3	176	4	US-09-252-991A-29512	Sequence 29512, A
43	39	44.3	195	4	US-09-252-991A-30082	Sequence 30082, A
44	39	44.3	204	4	US-09-252-991A-21317	Sequence 21317, A
45	39	44.3	468	4	US-09-252-991A-24394	Sequence 24394, A

ALIGNMENTS

```
RESULT 1
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-32326

Query Match      54.5%; Score 48; DB 4; Length 421;
Best Local Similarity 52.9%; Pred. No. 4.6;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

CY      2 HOESTXKXSRGSRGSG 18
db      13 HARTGTGRRGRRGSG 29

RESULT 2
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015

Query Match 50.0%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 H0ESTXGSRGSRGSGS 18
DB 94 HAECDFGAGTGRAGAG 110

RESULT 3

US-09-513-057C-13
Sequence 13, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 13
LENGTH: 653
TYPE: PRT
ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 50.0%; Score 44; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SH0ESTXGSRGSRGSGS 19
DB 573 SRQVSTASSAGSRGSGS 591

RESULT 4

US-08-591-502B-49
Sequence 49, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08665
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-591-502B-49

Query Match 50.0%; Score 44; DB 4; Length 730;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSJARGSRGSGS 242

RESULT 5

US-08-591-502B-50
Sequence 50, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08665
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-08-591-502B-50

Query Match 50.0%; Score 44; DB 4; Length 842;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGRSGS 242

RESULT 6
US-08-591-502B-51
Sequence 51, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51

Query Match 50.0%; Score 44; DB 4; Length 842;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGRSGS 242

RESULT 7
US-09-719-528A-2
Sequence 2, Application US/09719528A
Patent No. 6558675
GENERAL INFORMATION:

APPLICANT: Con, Chong Jin
Lim, Gek Keow
Zhao, Yi
Chen, Wei Ning
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Ladao & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2

Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGRSGS 242

RESULT 8
US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45
Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QESTXKSRGRSGSGS 19
Db 226 QPQGSMLRGKSGSGS 242
RESULT 9
US-08-591-502B-48
Sequence 48, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870

FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-08-591-502B-48
Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QESTXKSRGRSGSGS 19
Db 226 QPQGSMLRGKSGSGS 242
RESULT 10
US-08-591-502B-59
Sequence 59, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILRGKSGSGS 242

RESULT 11
US-08-591-502B-46
Sequence 46, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 845 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16..18
OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = unknown"
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-08-591-502B-46

Query Match 50.0%; Score 44; DB 4; Length 845;

Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGSRGSGS 19
DB 228 QPQGSILRGKSGSGS 244

RESULT 12
US-09-342-084-6
Sequence 6, Application US/09342084
Patent No. 6251668
GENERAL INFORMATION:
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Sakai, Hajime
TITLE OF INVENTION: Transcription Coactivators
FILE REFERENCE: BB-1169-A
CURRENT APPLICATION NUMBER: US/09/342,084
CURRENT FILING DATE: 1999-06-29
EARLIER APPLICATION NUMBER: 60/092,659
EARLIER FILING DATE: July 13, 1998
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: UNSURE
LOCATION: (128)
US-09-342-084-6

Query Match 48.9%; Score 43; DB 3; Length 169;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTXGSRGSRGSGS 19
DB 15 NNKXSGSGSRGSRPSGS 33

RESULT 13
US-09-252-991A-16961
Sequence 16961, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16961
LENGTH: 138
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16961

Query Match 47.7%; Score 42; DB 4; Length 138;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 8 GXSRGSRGSGS 19
DB 106 GPGRGGRGSGS 117

RESULT 14

```

US-09-252-991A-24514
; Sequence 24514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24514
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24514

```

```

Query Match          47.7%; Score 42; DB 4; Length 586;
Best Local Similarity 72.7%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      8 GXSGRSGRSG 18
      |||||
Db      252 GSGRSGRSG 262

```

```

RESULT 15
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4

```

```

Query Match          47.7%; Score 42; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      5 STYXGSRGSRSGS 19
      |||||
Db      562 SPGHSGRSGRSGS 576

```

```

Search completed: September 28, 2004, 06:26:44
Job time : 15.4479 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-7

Perfect score: 88
Sequence: 1 SHOESTXGRSXRGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	416	2 A32947	filaggrin precursor
2	85	96.6	2248	2 A35938	profilaggrin - hum
3	76	86.4	591	2 A45135	profilaggrin - hum
4	52	59.1	506	1 W2ML47	E2 protein - human
5	49	55.7	822	2 T51049	related to nucleol
6	48	54.5	471	2 T33997	hypothetical prote
7	47	53.4	306	2 T21220	hypothetical prote
8	47	53.4	399	2 T46259	hypothetical prote
9	44	50.0	123	2 T16234	hypothetical prote
10	44	50.0	800	2 T02852	probable membrane
11	43.5	49.4	286	2 S50855	neurotrophin-6 - s
12	43	48.9	649	2 G71283	probable ATP-depen
13	43	48.9	694	2 S71786	wingless receptor
14	43	48.9	849	2 A96592	hypothetical prote
15	43	48.9	1015	2 T42013	frequency clock pr
16	43	48.9	1507	2 B47328	natural killer cel
17	42	47.7	136	2 T35632	probable transposa
18	42	47.7	1119	2 T16720	hypothetical prote
19	42	47.7	1337	2 T30291	dextranase - strep
20	42	47.7	2420	2 A84652	hypothetical prote
21	41	46.6	625	2 A34615	profilaggrin - rat
22	41	46.6	745	2 G01025	serine/threonine p
23	41	46.6	774	2 I48609	probable serine/th
24	41	46.6	825	2 J04163	DNA-binding protei
25	41	46.6	836	2 G84727	probable DNA topoi
26	41	46.6	2233	2 S63347	acetyl-CoA carboxy
27	40	45.5	167	2 S05545	dehydrin 3 - maize
28	40	45.5	167	2 T52599	squamosa promoter
29	40	45.5	168	2 A39316	RAB-17 protein - m

30	40	45.5	168	2 S08633	RAB-17 protein - m
31	40	45.5	174	2 T52600	squamosa promoter
32	40	45.5	195	2 T37163	hypothetical prote
33	40	45.5	197	2 T18818	hypothetical prote
34	40	45.5	203	2 C87801	protein C10G11.9
35	40	45.5	312	2 A31846	130K paracrystall
36	40	45.5	315	2 S53589	SOL2 protein - yea
37	40	45.5	373	2 T02976	probable DNA bindi
38	40	45.5	454	2 A36643	protein kinase (EC
39	40	45.5	482	2 C86322	hypothetical prote
40	40	45.5	484	2 S53641	protein kinase c1k
41	40	45.5	508	2 T22440	hypothetical prote
42	40	45.5	526	2 A41698	cell division cont
43	40	45.5	638	2 I53169	cytokestatin 2 - hu
44	40	45.5	692	2 H71426	hypothetical prote
45	40	45.5	760	2 F75530	ribonuclease - Del

ALIGNMENTS

RESULT 1

A32947
filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of th
A/Reference number: A32947; MIMD:89286901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-references: GB:W24355; NID:9182604; PIDN:AA52454.1; PID:9182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1921-1921
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 96.6% Score 85; DB 2; Length 416;
Best Local Similarity 89.5% Pred. No. 4.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query 1 SHOESTXGRSXRGRSGS 19
Db 7 SHOESTXGRSXRGRSGS 25

RESULT 2
A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; MIMD:91064347; PMID:2248957
A/Accession: A35938
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:U02929
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1921-1921
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

A:Reference number: Z19393
A:Accession: T21220
A:Status: preliminary: translated from GB/EMBL/DBJ
submitted to the EMBL data library, March 1991

A.Molecule type: DNA
A.Residues: 1-306 <MIL>
A.Cross-references: EMBL:293379; PIDN:CAB07588.1; GSPDB:GN00023; CESP:F21H7.5
A.Experimental source: clone F21H7
C.Genetics:
A.Gene: CESP:F21H7.5
A.Map position: 5
A.Introns: 28/3; 146/3; 232/3
C.Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match
Best Local Similarity 53.4%; Score 47; DB 2; Length 306;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 SHQSTXGRSGRSG 18
DB 48 SHRSKTKRSGKSG 65

RESULT 8
T46259
Hypothetical protein DKFZp761E0323.1 - human (fragment)
C.Species: Homo sapiens (man)
C.Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
R.Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A.Reference number: Z23032
A.Accession: T46259
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-399 <AAA>
A.Cross-references: EMBL:R1137423
A.Experimental source: adult amygdala; clone DKFZp761E0323
C.Genetics:
A.Note: DKFZp761E0323.1

Query Match
Best Local Similarity 53.4%; Score 47; DB 2; Length 399;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHQSTXGRSGRSG 18
DB 336 SNRQRSGSGRSG 353

RESULT 9
T16234
Hypothetical protein F32A5.7 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C.Accession: T16234
R.Pauley, A.
submitted to the EMBL Data Library, July 1995
A.Description: The sequence of C. elegans cosmid F32A5.
A.Reference number: Z18482
A.Accession: T16234
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-123 <PAU>
A.Cross-references: EMBL:U20864; NID:g669026; PID:g669028; PIDN:AAC4661.1; CESP:F32A5.7
A.Experimental source: strain Bristol N2
C.Genetics:
A.Gene: CESP:F32A5.7
A.Introns: 2/3; 49/3; 75/3

Query Match
Best Local Similarity 50.0%; Score 44; DB 2; Length 123;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGRSGRSG 18
DB 93 RQSRGRGGGRG 108

RESULT 10
T02852
probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)
C.Species: Leishmania major
C.Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C.Accession: T02852; H81462
R.Wyler, P.J.
submitted to the EMBL Data Library, May 1998
A.Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
A.Reference number: Z14740
A.Accession: T02852
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-800 <MTL>
A.Cross-references: EMBL:AE001274; NID:g3264850; PID:g2266920
R.Wyler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A.Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A.Reference number: A81455; MUID:99178987; PMID:10077609
A.Accession: H81462
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-800 <PVL>
A.Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24675.1; PID:g2266920; GSPDB:GN001
A.Experimental source: strain MHOM/IL/81/Friedlin
C.Genetics:
A.Gene: L1439.4
A.Map position: 1

Query Match
Best Local Similarity 50.0%; Score 44; DB 2; Length 800;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 HQESTXGRSGRSG 19
DB 429 HRDGVRLSTTASGRSGS 446

RESULT 11
S50855
neurotrophin-6 - southern platyfish
C.Species: Xiphophorus maculatus (southern platyfish)
C.Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jan-2003
C.Accession: S50855
R.Goetz, R.; Koester, R.; Winkler, C.; Raulf, F.; Lottspeich, F.; Scharlt, M.; Thoenen, I
Nature 372, 266-269, 1994
A.Title: Neurotrophin-6 is a new member of the nerve growth factor family.
A.Reference number: S50855; MUID:95059452; PMID:7869471
A.Accession: S50855
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-286 <GOE>
A.Cross-references: GB:L36942; NID:g642673; PIDN:AAA61923.1; PID:g642674
C.Superfamily: nerve growth factor beta chain

Query Match
Best Local Similarity 49.4%; Score 43.5; DB 2; Length 286;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 1 SHQSTXGRSGRSG 19
DB 203 THRSGIVIGRSGRSG 224

RESULT 12
G71283
probable AMP-dependent RNA helicase - syphilis spirochete
C.Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C.Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 02-Feb-2001
C.Accession: G71283
R.Fraser, C.M.; Norris, S.D.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

erson, J.; Khajak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDev
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A/Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A/Reference number: A71250; MUID:98332770; PMID:9665876
A/Accession: G71283
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1649 <COL>
A/Cross-references: GB:AE001248; GB:AE000520; NID:g3323074; PIDN:AA065738.1; PID:g332307
A/Experimental source: strain Nichols
C/Genetics:
A/Gene: TP0770
C/Keywords: ATP; nucleotide binding; P-loop
F:91-98/Region: nucleotide-binding motif A (P-loop)
F:192-197/Region: nucleotide-binding motif B
F:196-199/Region: DEAD motif

Query Match 48.9%; Score 43; DB 2; Length 649;
Best Local Similarity 44.4%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQSTXGRSGRSGS 19
DB 364 HDSATYTHRYGRTGRAGS 381

RESULT 13

Wingless receptor precursor dfz2 - fruit fly (*Drosophila melanogaster*)
C/Species: *Drosophila melanogaster*
C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
C/Accession: S71786; S78444
R/Bhanot, P.; Brink, M.; Harryman Samos, C.; Heish, J.C.; Wang, Y.; Macke, J.P.; Andrew,
Nature 382, 225-230, 1996
A/Title: A new member of the frizzled family from *Drosophila* functions as a wingless rec
A/Reference number: S71786; MUID:96353971; PMID:8717036
A/Accession: S71786
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-694 <BHA>
A/Cross-references: EMBL:U65589
A/Note: mRNA was also sequenced
R/Bhanot, P.; Wang, Y.; Nathans, J.
submitted to the EMBL Data Library, July 1996
A/Reference number: S78444
A/Accession: S78444
A/Molecule type: DNA
A/Residues: 1-416; T, 418-694 <BHW>
A/Cross-references: EMBL:U65589; NID:g1518050; PIDN:AA047273.1; PID:g1518051
C/Genetics:
A/Gene: dfz2
A/Cross-references: FlyBase:FBgn0016797
C/Superfamily: fruit fly frizzled protein
C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 19
DB 182 SYTEAGSGSGSGSGSGS 200

RESULT 14

A96592
hypothetical protein F14C21.55 [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: A96592
R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: A96592
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-849 <STO>
A/Cross-references: GB:AE005173; NID:g11055759; PIDN:AA028231.1; GSPDB:GN00141
C/Genetics:
A/Gene: F14C21.55
A/Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGS 19
DB 736 SSANRSRGGRGIGS 750

RESULT 15

frequency clock protein - *Creopus spinulosus*
C/Species: *Creopus spinulosus*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C/Accession: T42013
R/Lewis, M.T.; Morgan, L.W.; Feldman, J.F.
Mol. Gen. Genet. 253, 401-414, 1997
A/Title: Analysis of frequency (frc) clock protein homologs: evidence for a helix-turn-he
A/Reference number: Z22024; MUID:9718515; PMID:9037100
A/Accession: T42013
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1015 <LEW>
A/Cross-references: EMBL:U25850; NID:g852501; PID:g852502; PIDN:AAA68072.1
C/Genetics:
A/Gene: frc
A/Introns: 100/1

Query Match 48.9%; Score 43; DB 2; Length 1015;
Best Local Similarity 42.1%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 19
DB 583 NHRKQKTGSHSGSGSGN 601

Search completed: September 28, 2004, 06:15:01
Job time : 13.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.352 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQSTYGRSXGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 141681 segs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	85	96.6	416	1	FILA_HUMAN
2	52	59.1	506	1	VE2_HPV47
3	44	50.0	123	1	LSM4_CAEEL
4	43	48.9	694	1	FRZ2_DROME
5	43	48.9	1015	1	FRO_CRESP
6	43	48.9	1453	1	NKCR_MOUSE
7	42	47.7	1337	1	DEXT_STRDO
8	41	46.6	774	1	MKX2_MOUSE
9	41	46.6	825	1	SES_RAT
10	41	46.6	2233	1	COAC_YEAST
11	40	45.5	168	1	DH1_MAIZE
12	40	45.5	315	1	SOL2_YEAST
13	40	45.5	484	1	CLK1_HUMAN
14	40	45.5	526	1	CLK1_YEAST
15	40	45.5	638	1	K220_HUMAN
16	40	45.5	1527	1	ARHB_RAT
17	39.5	44.9	135	1	RBP1_DROME
18	39	44.3	102	1	SAAS_MESAU
19	39	44.3	218	1	P25A_BOVIN
20	39	44.3	439	1	AP50_DICDI
21	39	44.3	831	1	FTSR_MYCTU
22	39	44.3	834	1	VN22_YEAST
23	38.5	43.8	857	1	Y2A_CWNT
24	38	43.2	327	1	FBR1_MOUSE
25	38	43.2	340	1	EBP2_MOUSE
26	38	43.2	370	1	ERWE_SACER
27	38	43.2	476	1	GAB3_CHICK
28	38	43.2	483	1	CLK1_MOUSE
29	38	43.2	512	1	FUS_BOVIN
30	38	43.2	518	1	FUS_MOUSE
31	38	43.2	526	1	FUS_HUMAN
32	38	43.2	636	1	DEDT_SCHPO
33	38	43.2	674	1	TML1_ARATH

34	38	43.2	706	1	NUCL_HUMAN	P13338	homo sapien
35	38	43.2	706	1	NUCL_MOUSE	P09405	mus musculus
36	38	43.2	712	1	NUCL_RAT	P13383	rattus norv
37	38	43.2	730	1	DPOL_HPV4	P12933	hepatitis b
38	38	43.2	842	1	DPOL_HPBVM	P18707	hepatitis b
39	38	43.2	843	1	DPOL_HPBVR	P03157	hepatitis b
40	38	43.2	1094	1	A3B1_HUMAN	O00203	homo sapien
41	37.5	42.6	716	1	DVJ3_HUMAN	O92997	homo sapien
42	37.5	42.6	716	1	DVJ3_MOUSE	O61062	mus musculus
43	37	42.0	122	1	SAI1_MOUSE	P05366	mus musculus
44	37	42.0	122	1	SAI1_MOUSE	P05367	mus musculus
45	37	42.0	325	1	FRZ2_BOVIN	O95117	bos taurus

ALIGNMENTS

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89296901; PubMed=2740331;
RA McKinnley-Grant L.U., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Camilizaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RN [2]
RP CITRULLINATION
RX MEDLINE=96374388; PubMed=8780679;
RA Senshu T., Kan S., Ogawa H., Manabe M., Aaga H.;
RT "Preferential deimination of keratin XI and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -I- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -I- PFM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -I- PFM: Undergoes deimination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M24355; AAA52454.1; -
CC PIR; A32947; A32947;
CC Genew; HGNC:3748; FIG.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.

DR Pfam: PF03516; Flagellin; 2.
 DR PRINTS: PR00487; FLAGGRIN.
 KM Phosphorylation; Citrullination; Developmental protein.
 FT NON_TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;
 Query Match 96.6%; Score 85; DB 1; Length 416;
 Best Local Similarity 89.5%; Pred. No. 1e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHOEYXGRSGRSGSGS 19
 Db 7 SHOEYXGRSGRSGSGS 25
 RESULT 2
 VE2_HP47 STANDARD; PRT; 506 AA.
 ID VE2_HP47
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT Genome organization and taxonomic position of human papillomavirus
 RT type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).
 CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE EPR RESPONSE ELEMENT (5'-ACCCNNNGGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2B'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M32305; AAA46979.1; -
 CC PIR: D35324; M2ML47.
 CC HSSP: P03122; 2BOP.
 CC InterPro: IPR000427; E2_C.
 CC InterPro: IPR001866; E2_N.
 CC InterPro: IPR009021; Viral_DNA_bd.
 CC Pfam: PF00511; E2_C; 1.
 CC Pfam: PF00508; E2_N; 1.
 CC ProDom: PD000672; E2_C; 1.
 CC ProDom: PD000678; E2_N; 1.
 KM Early protein: Transcription regulation; Activator; DNA-binding;
 KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF9725065 CRC64;
 Query Match 59.1%; Score 52; DB 1; Length 506;
 Best Local Similarity 52.6%; Pred. No. 0.33;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 SHOEYXGRSGRSGSGS 19

Db 342 SREGRGRGRGRGRGRGS 360
 RESULT 3
 LSM4_CAEEL STANDARD; PRT; 123 AA.
 ID LSM4_CAEEL
 AC Q19952;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable U6 snRNA-associated Sm-like protein LSM4.
 GN F32A5.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
 CC snRNA (By similarity).
 CC -1- SUBUNIT: Sm subunits form a heteromer with a doughnut shape (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U20864; AAC46661.1; -
 CC PIR: T16234; T16234.
 CC WormPep: F32A5.7; CE01277.
 CC InterPro: IPR006649; snRNP.
 CC InterPro: IPR001163; snRNP_Sm.
 CC Pfam: PF01423; LSM; 1.
 CC ProDom: PD020287; snRNP; 1.
 DR SMART: SM00651; Sm; 1.
 KM Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
 KM RNA-binding.
 SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;
 Query Match 50.0%; Score 44; DB 1; Length 123;
 Best Local Similarity 50.0%; Pred. No. 1.5;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 3 QESTYXGRSGRSGSG 18
 Db 93 QESTYXGRSGRSGRSG 108
 RESULT 4
 FRZ2_DROME STANDARD; PRT; 694 AA.
 ID FRZ2_DROME
 AC Q9VFX3; Q94916; Q9VFX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frizzled protein 2 precursor (Frizzled-2) (DFZ2).
 GN FRZ OR CG9739/CG14083.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
RX MEDLINE=96353971; PubMed89717036;
RA Bhanoji P., Brink M., Santos C.H., Hsieh J.C., Wang Y., Macke J.P.,
RA Andrew D., Nathans J., Nüssle R.,
RT "A new member of the Fz/zled family from *Drosophila* functions as a
RT Wingless receptor.";
RL Nature 382:225-230(1996).

RP SEQUENCE FROM N.A.

FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors are coupled to the beta-catenin canonical signaling pathway, which leads to the activation of dishevelled proteins, inhibition of GSK-3 kinase, nuclear accumulation of beta-catenin and activation of Wnt target genes. A second signaling pathway involving PKC and calcium fluxes has been seen for some family members, but it is not yet clear if it represents a distinct pathway or if it can be integrated in the canonical pathway, as PKC seems to be required for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem to involve interactions with G-proteins. Required to coordinate the cytoskeletons of epidermal cells to produce a parallel array of cuticular hairs and bristles.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
CC between 15 and 70 per cent of egg length, including the
CC invaginating cells of the ventral furrow. Stripe pattern is
CC emerging by early stage 8. From stage 9 and continuing throughout
CC embryogenesis, expression is seen in the developing CNS. At stage
CC 10, expressed in 15 stripes in the presumptive head and trunk
CC regions, in the posterior midgut primordium, in a subset of cells
CC of anterior midgut invagination and in the procephalic lobe. At
CC stage 12, expression declines in epidermis and increases in the
CC midgut and visceral mesoderm. At stage 17, only expressed in the
CC CNS, hindgut and dorsal vessel.

--
CC - DOMAIN: Lys-Thr-X-X-Thr motif is involved in the activation
CC of the Wnt/Beta-catenin signaling pathway (By similarity).
CC - DOMAIN: The Fz domain is involved in binding with Wnt ligands.
CC - SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
CC family.
CC - SIMILARITY: Contains 1 frizzled (FZ) domain.
CC
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/announcement/>
CC or send an email to license@ebi.ac.uk).
CC

CC
DR EMBL; U65589; AAC47273.1; -.
DR EMBL; AE003518; AAF49185.2; -.
DR PIR; S71786; S71786.
DR FlyBase; FBgn0016797; f22.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0017147; F:Mnt-protein binding; IDA.
DR GO; GO:0006328; P:cell motility; IMP.
DR GO; GO:0007163; P:establishment and/or maintenance of cell po. .; ISS.
DR GO; GO:0008585; P:female gonad development; IMP.
DR GO; GO:0016055; P:Mnt receptor signaling pathway; IDA.
DR InterPro; IPR000539; Fz1zled.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR000832; GPCR_secretin.
DR Pfam; PF01534; Fz1zled; 1.

[illegible]

```

FRQ_CRESP          STANDARD;          PRT; 1015 AA.
ID   FRQ_CRESP
AC   Q00586;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Frequency clock protein.
GN   FRQ.
OS   Crepus spullosus (Chromocera spullosa).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX   NCBI_TaxId=110619;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FGSC 4896;
RX   MEDLINE=97186515; PubMed=90371100;
RA   Lewis M.T., Morgan L.W., Feldman J.F.;
RT   "Analysis of frequency (frq) clock gene homologs: evidence for a
RT   helix-turn-helix transcription factor.";
RL   Mol. Gen. Genet. 253:401-414(1997).
CC   -!- FUNCTION: Circadian clock component involved in the generation of
CC   biological rhythms, in particular in rhythm stability, period
CC   length, and temperature compensation. Behaves as a negative
CC   element in circadian transcriptional loop (by similarity).
CC   -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; U25850; AAA66072.1; -.
DR   F1R; T42013; T42013.
KW   Biological rhythms; Transcription regulation; Nuclear protein.
FT   DOMAIN          240..245
FT   DOMAIN          356..368
FT   DOMAIN          443..451
FT   DOMAIN          584..588
FT   DOMAIN          892..913
FT   DOMAIN          1015 AA; 110972 MW; EA49E732BD7A14B1 CRC64;
SQ   SEQUENCE          1015 AA; 110972 MW; EA49E732BD7A14B1 CRC64;

Query Match          48.9%; Score 43; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY   1 SHQSTXGRSGXGRSGS 19
DB   583 NHRKQXGTGHTGSDSGSSGN 601

RESULT 6
NCR_MOUSE          STANDARD;          PRT; 1453 AA.
ID   NCR_MOUSE
AC   P30415;
DT   01-APR-1993 (Rel. 25, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   NK-tumor recognition protein (Natural-killer cells cyclophilin-
DE   related protein) (NK-TR protein).
GN   NKTR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxId=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9313824; PubMed=8421688;
RA   Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
RA   Ortaldo J.R.;
RT   "A cyclophilin-related protein involved in the function of natural

```

```

RT   killer cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RN   [2]
RP   REVISIONS TO C-TERMINUS.
RC   STRAIN=BALB/c; TISSUE=Blood;
RA   Anderson S.K.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DBD databases.
CC   -!- FUNCTION: Component of a putative tumor-recognition complex.
CC   Involved in the function of NK cells.
CC   -!- SIMILARITY: Contains 1 cyclophilin-like PPIase domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; L04289; AAA37500.2; ALT_INIT.
DR   HSSP; Q27450; IAK3.
DR   MGJ; MGJ:97346; NKtr.
DR   InterPro; IPR002130; CSA_PPIase.
DR   Pfam; PF00160; PRO_Isomerase_1.
DR   PRINTS; PR00153; CSAPISMRASE.
DR   PROSITE; PS00170; CSA_PPIASE_1; 1.
DR   PROSITE; PS50072; CSA_PPIASE_2; 1.
KW   Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
FT   DOMAIN          1..176
FT   DOMAIN          222..241
FT   DOMAIN          422..459
FT   DOMAIN          964..1003
FT   DOMAIN          198..273
FT   DOMAIN          468..565
FT   DOMAIN          658..812
FT   DOMAIN          1303..1453
FT   DOMAIN          163439 MW; DFL173FFB14B283B CRC64;
SQ   SEQUENCE          1453 AA; 163439 MW; DFL173FFB14B283B CRC64;

Query Match          48.9%; Score 43; DB 1; Length 1453;
Best Local Similarity 81.8%; Pred. No. 30;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY   9 RSKGRSGRSGS 19
DB   682 RSGSGSGRSGS 692

RESULT 7
DEXT_STRDO          STANDARD;          PRT; 1337 AA.
ID   DEXT_STRDO
AC   P39653;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
DE   glucanohydrolase).
GN   DEX.
OS   Streptococcus downei (Streptococcus sobrinus).
OG   Plasmid PYA902.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxId=13117;
RN   [1]
RP   SEQUENCE FROM N.A. AND SEQUENCE OF 31-36.
RC   STRAIN=6715 / UAB65;
RX   MEDLINE=94292401; PubMed=8021165;
RA   Wanda S.-Y., Curtiss R. III;
RT   "Purification and characterization of Streptococcus sobrinus
RT   dextranase produced in recombinant Escherichia coli and sequence
RT   analysis of the dextranase gene.";
RL   J. Bacteriol. 176:3839-3850(1994).
CC   -!- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
CC   PELLICLE-COATED TOOTH SURFACE.

```

```

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
CC linkages in dextran.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
CC and at 39 degrees Celsius.
CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M96978; AAA21772.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfam; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
KM Signal; Plasmid.
FT SIGNAL 1 30
FT CHAIN 1 1308 DEXTRANSASE.
FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 17
DB 1154 ANQDSTKSGSADQSGKS 1170

RESULT 8
MRK2_MOUSE STANDARD; PRT; 774 AA.
ID MRK2_MOUSE
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
DE MARK2 OR EMK.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=9336412; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19."
RL Mamm. Genome 4:401-403 (1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70764; CA50040.1; -.
DR PIR; I48609; I48609.
DR HSP; Q63450; I406.
DR MGD; MGI:99638; Mark2.
DR InterPro; IPR001772; Kinase_Cterm.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Enr_kin_AS.
DR InterPro; IPR002290; Ser_Chr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; PKase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 53 304
FT NP_BIND 59 67 PROTEIN KINASE.
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

Query Match 46.6%; Score 41; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 STYGRSXGRSGRSGS 19
DB 614 SPFGHSQGRGRGASGS 628

RESULT 9
S55_RAT STANDARD; PRT; 825 AA.
ID S55_RAT
AC Q63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE S55 antigen.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=9601519; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein S55 in the nervous system."
RL J. Biochem. 118:122-128 (1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D37934; BA07153.1; -.
DR PIR; J04163; J04163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86631 MW; AF667FE2FD55BDF CRC64;

Query Match 46.6%; Score 41; DB 1; Length 825;

```

Best Local Similarity 66.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CY 8 GRSXGRSGRGS 19
DB 589 GRSXGRSGRGS 600

RESULT 10
COAC YEAST STANDARD; PRT: 2233 AA.
AC 000955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin
carboxylase (EC 6.3.4.14)].
GN FAS3 OR ACC1 OR YNR016C OR N3175.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2015-2022.
RX MEDLINE=92262474; PubMed=1350093;
RA Al-Feel W., Chirala S.S., Wakil S.J.;
RT "Cloning of the yeast FAS3 gene and primary structure of yeast
acetyl-CoA carboxylase."
RT Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538 (1992).
RL [2]
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This protein carries three functions: biotin carboxyl
carrier protein, biotin carboxylase, and carboxyltransferase.
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
+ malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
= ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: Biotin.
CC -1- ENZYME REGULATION: By phosphorylation.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M2156; AAA20073.1; -;
DR EMBL, Z71631; CA96294.1; -;
DR PIR, S63347; S63347.
DR HSSP, P24182; 1BNC.
DR GenOnline, 14361; -;
DR SGD, S0005299; ACC1.
DR GO, GO:0005789; C:cytoplasmic reticulum membrane; IDA.
DR GO, GO:0003989; F:acetyl-CoA carboxylase activity; IMP.
DR GO, GO:0004075; F:biotin carboxylase activity; IMP.
DR InterPro, IPR001882; Biotin_BS.
DR InterPro, IPR005482; Biotin_carb_C.
DR InterPro, IPR000089; Biotin_lipoyl.
DR InterPro, IPR000022; Carboxyl_trans.
DR InterPro, IPR005479; Cphase_L_D2.
DR InterPro, IPR005481; Cphase_L_N.
DR Pfam, PF02785; Biotin_carb_C_1.
DR Pfam, PF00364; Biotin_lipoyl_1.
DR Pfam, PF01039; Carboxyl_trans_1.

DR Pfam, PF00289; CPSase_L_chain; 1.
DR Pfam, PF02786; CPSase_L_D2; 1.
DR PROSITE, PS00188; BIOTIN; 1.
DR PROSITE, PS00866; CPSASE_1; 1.
DR PROSITE, PS00867; CPSASE_2; 1.
KM Fatty acid biosynthesis; biotin; ligase; Multifunctional enzyme;
KM ATP-binding; phosphorylation.
FT NP_BIND 256 261 ATP (BY SIMILARITY).
FT ACT_SITE 383 383 BY SIMILARITY.
FT BINDING 735 735 BIOTIN (BY SIMILARITY).
FT DOMAIN 1865 1894 COENZYME A-BINDING (BY SIMILARITY).
FT CONFLICT 1523 1523 W -> G (IN REF. 1).
FT CONFLICT 1755 1755 I -> TWYRCL (IN REF. 1).
FT CONFLICT 1761 1766 AINRML -> ESTNA (IN REF. 1).
SQ SEQUENCE 2233 AA; 250351 MW; 0A355AD9B1F8308 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 2233;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
CY 2 HESTXGRSGRSGRGS 19
DB 1200 HSSXNGPAPPGRSSAS 1217

RESULT 11
DH1_MAIZE STANDARD; PRT: 168 AA.
AC P12950; P16318;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Dehydrin DHN1 (M3) (RAB-17 protein).
GN DHN1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93357436; PubMed=2562763;
RA Close T.J., Kott A.A., Chandler P.M.;
RT "A cDNA-based comparison of dehydration-induced proteins (dehydrins)
in barley and corn."
RT Plant Mol. Biol. 13:95-108 (1989).
RL [2]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, AC 1503;
RX MEDLINE=91346629; PubMed=2151715;
RA Villardell J., Goday A., Freire M.A., Torrent M., Martinez M.C.,
Torre J.M., Pages M.;
RT "Gene sequence, developmental expression, and protein phosphorylation
of RAB-17 in maize."
RT Plant Mol. Biol. 14:423-432 (1990).
RN [3]
RP PHOSPHORYLATION BY CKII.
RX MEDLINE=92042196; PubMed=1939268;
RA Plana M., Iarte E., Ritzja R., Goday A., Pages M., Martinez M.C.;
RT "Phosphorylation of maize RAB-17 protein by casein kinase 2."
RL J. Biol. Chem. 266:22510-22514 (1991).
CC -1- DEVELOPMENTAL STAGE: ABA treatment induced the synthesis of RAB-17
in calli, however, the RAB-17 proteins were found to be highly
phosphorylated only in embryos.
CC -1- INDUCTION: By abscisic acid (ABA) and water stress.
CC -1- PTM: SERINE IS THE ONLY AMINO ACID WHICH BECOMES PHOSPHORYLATED.
CC -1- SIMILARITY: Belongs to the plant dehydrin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL, X15290; CAA33364.1; -
 DR EMBL, X15994; CAA34123.1; -
 DR PIR, A39316; A39316.
 DR PIR, S05545; S05545.
 DR PIR, S08633; S08633.
 DR MatzedeB, 24916; -
 DR InterPro, IPR00167; Dehydrin.
 DR Pfam, PF00257; dehydrin.1.
 DR ProSite, PS00315; DEHYDRIN_1; 1.
 DR ProSite, PS00823; DEHYDRIN_2; 2.
 DR Dehydrin; Repeat; Phosphorylation; Multigene family.
 KW DOMAIN 76 84 POLY-SER
 FT DOMAIN 2 140 3 X APPROXIMATE REPEATS.
 FT REPEAT 2 10 1-1.
 FT REPEAT 122 130 1-1.
 FT REPEAT 132 140 1-2.
 FT DOMAIN 94 166 2 X APPROXIMATE REPEATS.
 FT REPEAT 94 107 2-1.
 FT REPEAT 153 166 2-2.
 FT CONFLICT 9 9 R -> H (IN REF. 2).
 FT CONFLICT 13 13 R -> A (IN REF. 2).
 FT CONFLICT 41 41 MISSING (IN REF. 2).
 SQ SEQUENCE 168 AA; 17161 MW; BE7758CAD37CF39 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 168;
 Best Local Similarity 38.9%; Pred. No. 9.3;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 HOESTXGRSGRSGSGS 19
 DB 114 HATATTGGAGYGGGHTGS 131

RESULT 12
 SOL2_YEAST STANDARD; PRT; 315 AA.
 AC P37262;
 DT 01-OCT-1996 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (6PGL).
 GN SOL2 OR YCRX13W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96363906; PubMed=8725220;
 RA Shen W.C., Stanford D.R., Hopper A.K.;
 RT "Loss of, involved in yeast pre-rRNA splicing, positively regulates
 RT members of the SOL gene family.";
 RL Genetics 143:699-712(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ballesta U.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
 RA Sanz E.;
 RT Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION AND SIMILARITY.
 RX MEDLINE=9414796; PubMed=831894;
 RA Koonin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RL EMBL J. 13:493-503(1994).
 CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
 CC phosphogluconate. (By similarity).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
 CC phospho-D-gluconate.
 CC -1- PATHWAY: Pentose phosphate pathway; second step.
 CC -1- SIMILARITY: Belongs to the glucosamine-6-phosphate

CC isomerase family. 6-phosphogluconolactonase subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL, U46559; AAB49322.1; -
 DR EMBL, X59720; CAA42272.1; -
 DR PIR, S53589; S53589.
 DR Germonline, 139004; -
 DR SGD, S0000718; SOL2.
 DR GO, GO:0008033; P:RNA processing; IGI.
 DR InterPro, IPR006146; Gluc_gal_isom.
 DR InterPro, IPR005900; Phosphogluconlac.
 DR Pfam, PF01182; Glucosamine_iso; 1.
 DR TIGRFAMs, TIGR01198; pgl, 1.
 KW Hydrolase.
 FT CONFLICT 171 171 A -> V (IN REF. 2).
 SQ SEQUENCE 315 AA; 34501 MW; 7F0AAD7657AB276 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 315;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 ESTXGRSGRSGSGS 19
 DB 60 KSTASAEGRSGSGS 75

RESULT 13
 CLTK1_HUMAN STANDARD; PRT; 484 AA.
 AC P49759;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein kinase CLK1 (EC 2.7.1.37) (EC 2.7.1.112) (CDC
 DE like kinase 1).
 GN CLTK1 OR CLK.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91138618; PubMed=1704859;
 RA Johnson K.W., Smith K.A.;
 RT "Molecular cloning of a novel human cdc2/CDC28-like protein kinase.";
 RL J. Biol. Chem. 266:3402-3407(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95082033; PubMed=7990150;
 RA Hanes J.U., der Kammer H., Klaudiny J.J., Scheit K.H.;
 RT "Characterization by cDNA cloning of two new human protein kinases.
 RT Evidence by sequence comparison of a new family of mammalian protein
 RT kinases.";
 RL J. Mol. Biol. 244:665-672(1994).
 CC -1- FUNCTION: PHOSPHORYLATES SERINE- AND ARGinine-RICH (SR) PROTEINS
 CC OF THE SPliceosomal COMPLEX MAY BE A CONSTITUENT OF A NETWORK OF
 CC REGULATORY MECHANISMS THAT ENABLE SR PROTEINS TO CONTROL RNA
 CC SPLICING. PHOSPHORYLATES SERINES, THREONINES AND TYROSINES (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: Named isoforms=2;
 CC Name=Long;

CC IsoId=P49759-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P49759-2; Sequence=VSP_004852, VSP_004853;
CC Note=Lacks the kinase domain;
CC -1- PIM: Autophosphorylates on all three types of residues (By
CC similarity). Belongs to the Ser/Thr family of protein kinases.
CC Lammer subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; I29219; AAA61480.1; -
DR EMBL; I29222; AAB59459.1; -
DR PIR; S53641; S53641.
DR HSSP; Q00534; 1B18.
DR Genew; HGNC:2068; CLK1.
DR GK; P49759; -
DR MM; 601951; -
DR GO; GO:0004715; P:non-membrane spanning protein tyrosine kina. ; TAS.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR008271; Ser_Thr_Pkin_A5.
DR InterPro; IPR002290; Ser_Thr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR TranserBase; Serine/threonine-protein kinase; ATP-binding;
KM Tyrosine-protein kinase; Phosphorylation; Nuclear protein;
KM Alternative splicing.
FT DOMAIN 161 477 PROTEIN KINASE.
FT NP_BIND 167 175 ATP (BY SIMILARITY).
FT BINDING 191 191 ATP (BY SIMILARITY).
FT ACT_SITE 288 288 BY SIMILARITY.
FT VARSPIC 131 136 KSHRRK -> MKLLIL (in isoform Short).
FT VARSPIC 137 484 /FTId=VSP_004852.
FT Missing (in isoform Short).
FT /FTId=VSP_004853.
SQ SEQUENCE 484 AA; 57205 MW; 304B5B486AD0A6B CRC64;
Query Match 45.5%; Score 40; DB 1; Length 484;
Best Local Similarity 47.1%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
CY 1 SHOESTXGRSGRSGRS 17
DB 95 SRYQNHSSKSGRSGRS 111
RESULT 14
ZABA YEAST
ID ZABA YEAST STANDARD; PRT; 526 AA.
AC Q00362;
DI 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein phosphatase PP2A regulatory subunit B (PR55) (Cell division
DE control protein 55).
DE CDC55 OR YGL190C OR G1345.
GN Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92017858; PubMed=1656238;
RA Healy A.M., Zolnerowicz S., Stapleton A.E., Goebel M.,
RA Depoli-Rocha A.A., Pringle J.R.;
RT "CDC55, a Saccharomyces cerevisiae gene involved in cellular
RT morphogenesis: identification, characterization, and homology to the
RT B subunit of mammalian type 2A protein phosphatase";
RT Mol. Cell. Biol. 11:5767-5780(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / FY1679;
RX MEDLINE=97197971; PubMed=9046087;
RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA Brusch C.V.;
RT "Sequencing of a 40.5 kb fragment located on the left arm of
RT chromosome VII from Saccharomyces cerevisiae";
RT Yeast 13:55-64(1997).
CC -1- FUNCTION: Phosphatase 2A affects a variety of biological processes
CC in the cell such as transcription, cell cycle progression and
CC cellular morphogenesis, and provides an initial identification of
CC critical substrates for this phosphatase. The regulatory subunit
CC may direct the catalytic subunit to distinct, albeit overlapping,
CC subsets of substrates.
CC -1- SUBUNIT: PP2A exists in several trimeric forms, all of which
CC consist of a core composed of a catalytic subunit associated with
CC a 65 kDa (PR65) (Subunit A) and a 55 kDa (PR55) (Subunit B)
CC regulatory subunit.
CC -1- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
CC family.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M72716; AAA34482.1; -
DR EMBL; Z72712; CAA96902.1; -
DR EMBL; X91837; CAA62954.1; -
DR EMBL; X91489; CAA62785.1; -
DR PIR; A41698; A41698.
DR Germonline; 141238; -
DR SGD; S0003158; CDC55.
DR GO; GO:0007094; P:mitotic spindle checkpoint; IMP.
DR GO; GO:0006412; P:protein biosynthesis; IMP.
DR GO; GO:0007124; P:pseudohyphal growth; IMP.
DR InterPro; IPR000009; PP2A_PR55.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR PRINTS; PR00600; PP2APR55.
DR SMART; SMO0320; WD40; 3.
DR PROSITE; PS01024; PR55_1; 1.
DR PROSITE; PS01025; PR55_2; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; FALSE NEG.
DR PROSITE; PS50294; WD_REPEATS_REGION; FALSE NEG.
KW Repeat: WD repeat; Cell cycle.
FT REPEAT 82 123 WD 1.
FT REPEAT 82 123 WD 1.
FT REPEAT 182 220 WD 2.
FT REPEAT 231 271 WD 4.
FT REPEAT 290 328 WD 5.
FT REPEAT 345 386 WD 6.
FT REPEAT 495 525 WD 7.
FT DOMAIN 416 419 POLY-SER.
FT CONFLICT 500 500 I -> N (IN REF. 1).
SQ SEQUENCE 526 AA; 59662 MW; 6DA12C2805FA6A82 CRC64;
Query Match 45.5%; Score 40; DB 1; Length 526;

Best Local Similarity 42.1%; Pred. No. 31;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
OY 1 SHOESTGRSXGRSGRS 19
DB 422 SHKRSNGRNTGWSSNS 440

RESULT 15

K220 HUMAN STANDARD: PRT: 638 AA.

AC 001546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (CytoKeratin 2P) (K2P)
DE (CK 2P).
GN KRT2P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93114504; PubMed=1282112;
RA Collin C., Ouhayoun J.P., Grunz C., Franke W.W.;
RT "Suprabasal marker proteins distinguishing keratinizing squamous
RT epithelia: cytoKeratin 2 polypeptides of oral masticatory epithelium
RT and epidermis are different."
RL Differentiation 51:137-148(1992).
CC -1- FUNCTION: Probably contributes to terminal cornification.
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC keratinocytes.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin: I (acidic; 40-55 kDa) [K1 to K9] and II
CC (neutral to basic; 56-70 kDa) [K1 to K9].
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M99063; AAA5746.1; -
DR PIR: I53169; I53169.
DR GO: GO:0005882; C:intermediate filament; NAS.
DR GO: GO:0005198; P:structural molecule activity; NAS.
DR GO: GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
DR InterPro: IPR001664; IF.
DR InterPro: IPR002957; Keratin_I.
DR InterPro: IPR003054; Keratin_II.
DR Pfam: PF00038; filament; 1.
DR PRINTS: PRO1248; TYPE1KERATIN.
DR PRINTS: PRO1276; TYPE2KERATIN.
DR PROSITE: PS00226; IF_1.
KW Intermediate filament; Coiled coil; Keratin; Phosphorylation.
FT DOMAIN 1 182 HEAD.
FT DOMAIN 183 452 ROD.
FT DOMAIN 453 638 TAIL.
FT DOMAIN 183 218 COIL_1A.
FT DOMAIN 219 237 LINKER_1.
FT DOMAIN 238 329 COIL_1B.
FT DOMAIN 330 353 LINKER_12.
FT DOMAIN 354 452 COIL_2.
FT MOD_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 638 AA: 65871 MW: 98743AB872076AF CRC64;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 3 QESTXGRSXGRSGRS 17
DB 7 KKSFGSRSGQFGSGRS 21

Search completed: September 28, 2004, 06:05:51
Job time: 8.22396 secs

Query Match 45.5%; Score 40; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 38;

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-7

Perfect score: 88

Sequence: 1 SHOESTYGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	85	96.6	798	4 Q9H4U3	Q9H4U3 homo sapien
2	85	96.6	1084	4 Q01212	Q01212 homo sapien
3	84	95.5	797	4 Q16824	Q16824 homo sapien
4	84	95.5	990	4 Q15206	Q15206 homo sapien
5	84	95.5	1218	4 Q05531	Q05531 homo sapien
6	76	86.4	591	4 Q01720	Q01720 homo sapien
7	76	86.4	687	4 Q9H4U2	Q9H4U2 homo sapien
8	71	80.7	465	4 Q03838	Q03838 homo sapien
9	70	79.5	322	4 Q75370	Q75370 homo sapien
10	51	58.0	1284	5 Q9V9Y3	Q9V9Y3 drosophila
11	49	55.7	822	3 Q9P312	Q9P312 neurospora
12	48	54.5	471	5 Q9UAY0	Q9UAY0 caenorhabdi
13	47	53.4	306	5 Q45386	Q45386 caenorhabdi
14	47	53.4	399	4 Q9NTA9	Q9NTA9 homo sapien
15	47	53.4	436	16 Q82C67	Q82C67 streptomyce
16	47	53.4	467	4 Q9BW18	Q9BW18 homo sapien

17	47	53.4	737	4 Q9BQ39	Q9BQ39 homo sapien
18	46	52.3	233	10 Q9AX45	Q9AX45 cryza sariv
19	46	52.3	426	4 Q8WV76	Q8WV76 homo sapien
20	45	51.1	259	4 Q86YA2	Q86YA2 homo sapien
21	45	51.1	669	11 Q8BRU5	Q8BRU5 mus musculu
22	45	51.1	724	11 Q99M79	Q99M79 mus musculu
23	44	50.0	216	4 Q9H711	Q9H711 homo sapien
24	44	50.0	322	16 Q8G4F5	Q8G4F5 bifidobacte
25	44	50.0	800	5 Q15845	Q15845 leishmania
26	44	50.0	936	16 Q82H71	Q82H71 streptomyce
27	43.5	49.4	286	13 Q91988	Q91988 xiphophorus
28	43	48.9	436	16 Q9AX30	Q9AX30 streptomyce
29	43	48.9	649	16 Q83749	Q83749 treporema p
30	43	48.9	813	12 Q91IN8	Q91IN8 hepatitis b
31	43	48.9	820	12 Q68375	Q68375 mouse cytom
32	43	48.9	869	10 Q9C7M2	Q9C7M2 arabidopsis
33	43	48.9	849	10 Q949M6	Q949M6 arabidopsis
34	43	48.9	966	11 Q8CBP6	Q8CBP6 mus musculu
35	43	48.9	1064	3 Q96X23	Q96X23 yarrowia li
36	43	48.9	1142	13 Q7S2V2	Q7S2V2 xenopus lae
37	42.5	48.3	240	16 Q87YL4	Q87YL4 pseudomonas
38	42	47.7	297	6 Q8SQ24	Q8SQ24 bos taurus
39	42	47.7	481	16 Q7UDP9	Q7UDP9 rhodospirill
40	42	47.7	552	11 Q8K028	Q8K028 mus musculu
41	42	47.7	658	11 Q8C5U9	Q8C5U9 mus musculu
42	42	47.7	806	12 Q91522	Q91522 hepatitis b
43	42	47.7	809	5 Q960C5	Q960C5 drosophila
44	42	47.7	1135	5 Q9VJB6	Q9VJB6 drosophila
45	42	47.7	1229	3 Q86ZM1	Q86ZM1 neurospora

ALIGNMENTS

RESULT 1					
Q9H4U3		PRELIMINARY;	PRT;	798 AA.	
AC Q9H4U3;					
DT 01-MAR-2001 (TRENBLREL.16, Created)					
DT 01-MAR-2001 (TRENBLREL.16, Last sequence update)					
DT 01-JUN-2003 (TRENBLREL.24, Last annotation update)					
DE DJI4N1.1.2 (Profilaggrin 3' end) (Fragment).					
GN FLG.					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OX NCBI_Taxid=9606;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Latid G.;					
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AL356504; CAC13171.1; -					
DR GO; GO:0005198; F:structural molecule activity; IEA.					
DR InterPro; IPR003303; Filaggrin.					
DR Pfam; PF03516; Filaggrin; 4.					
DR PRINTS; PRO0487; FILAGGRIN.					
FT NON_TER					
FT SEQUENCE	798 AA;	84773 MW;	P923DDABD1290805 CRC64;		
Query Match	96.6%;	Score 85;	DB 4;	Length 798;	
Best Local Similarity	89.5%;	Pred. No. 5e-06;			
Matches	17;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
QY	1 SHOESTYGRSGRSGSGS 19				
DB	428 SHOESTYGRSGRSGSGS 446				
RESULT 2					
Q01212		PRELIMINARY;	PRT;	1084 AA.	
ID Q01212					
AC Q01212; Q03840;					
DT 01-NOV-1996 (TRENBLREL.01, Created)					

```
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER
SQ SEQUENCE 1084 AA; 115271 MW; 80046408BD5A362D CRC64;

Query Match 96.5%; Score 85; DB 4; Length 1084;
Best Local Similarity 89.5%; Pred. No. 7e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTYGRSGRSGRSGS 19
Db 64 SHOESTGRSGRSGRSGS 82

RESULT 3
Q16824 PRELIMINARY; PRT; 797 AA.
ID Q16824;
AC Q16824;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60502; AAA63248.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BD86B CRC64;

Query Match 95.5%; Score 84; DB 4; Length 797;
Best Local Similarity 89.5%; Pred. No. 7.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTYGRSGRSGRSGS 19
Db 427 SHOESTGRSGRSGRSGS 445
```

```
RESULT 4
Q15206 PRELIMINARY; PRT; 990 AA.
ID Q15206;
AC Q15206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60494; AAA63244.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON_TER
SQ SEQUENCE 990 AA; 106453 MW; A8396F10FA931991 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 9.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTYGRSGRSGRSGS 19
Db 875 SHOESTGRSGRSGRSGS 893

RESULT 5
Q05331 PRELIMINARY; PRT; 1218 AA.
ID Q05331;
AC Q05331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS DURING
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
```

CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFLAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1; -
DR PIR; A48118; A48118.
DR HSP; P02593; 1CDM.
DR GO; GO:0005856; F:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; F:cell differentiation; NAS.
DR GO; GO:0008151; F:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Flaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FLAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR KX Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 SHQESTXGRSGRSGSGS 19
DB 1097 SHQESTXGRSGRSGSGS 1115

RESULT 6
ID Q01720 PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FLAGGRIN precursor (PROFLAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Nitsunskii W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus";
RL J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFLAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -
DR EMBL; L01090; AAA60176.1; -

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSP; P80511; 1E8A.
DR MIM; 135940; -
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005398; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Flaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FLAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293 POTENTIAL.
FT CHAIN 294 467 FLAGGRIN.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FLAGGRIN.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 86.4%; Score 76; DB 4; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHQESTXGRSGRSGSGS 19
DB 449 SHQESTXGRSGRSGSGS 467

RESULT 7
ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -
DR PIR; A48118; A48118.
DR HSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Flaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FLAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 86.4%; Score 76; DB 4; Length 687;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY      1  SHOESTYGRSGRSGSGS 19
      ||||| ||||| |||||
DB      449 SHOESTYGRSGRSGSGS 467

RESULT 8
ID      003838      PRELIMINARY;      PRT;      465 AA.
AC      003838:
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      FILAGGRIN (PROFILAGGRIN) (Fragment).
GN      FLG.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX      NCBL_TaxID=9606;
      (1)
RN      SEQUENCE FROM N.A.
RP      TISSUE=PLACENTA;
RC      MEDLINE=91064347; PubMed=2248957;
RX      MEDLINE=91064347; PubMed=2248957;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinart P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 29:9432-9440(1990).
      [2]
RN      REVISIONS.
RX      MEDLINE=91255199; PubMed=2043621;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinart P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 30:5814-5814(1991).
CC      -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
      PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
      FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC      -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
      WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
      REPEATS.
CC      -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
      INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDDEM
      COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN
      GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
      PROTEOLYTICALLY CLEAVED.
DR      EMBL, M60499; AAA63246.1; -.
DR      GO: GO:0005198; F:structural molecule activity; IEA.
DR      InterPro: IPR003303; Filaggrin.
DR      Pfam: PF03516; Filaggrin; 3.
DR      PRINTS: PR00487; FILAGGRIN.
FT      NON_TER      1
FT      NON_TER      1
SQ      SEQUENCE 465 AA; 50280 MW; C883744C5B134097 CRC64;

Query Match      80.7%; Score 71; DB 4; Length 465;
Best Local Similarity 78.9%; Pred. No. 0.00069;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY      1  SHOESTYGRSGRSGSGS 19
      ||||| ||||| |||||
DB      227 SHOESTYGRSGRSGSGS 245

RESULT 9
ID      075370      PRELIMINARY;      PRT;      322 AA.
AC      075370:
DT      01-NOV-1998 (TREMBlrel. 08, Created)
DT      01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT      01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE      Bidermal filaggrin (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

OX      NCBL_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99101527; PubMed=9886436;
RA      Girbal-Neuhauser E., Durieux J.J., Arnaud M., Dalbon P., Sabbag M.,
RA      Vincent C., Simon M., Senshu T., Masson-Bessiere C.,
RA      Jolivet-Reynaud C., Jolivet M., Serre G.;
RT      "The epitopes targeted by the rheumatoid arthritis-associated
      anti-filaggrin autoantibodies are posttranslational generated on
      various sites of (pro)filaggrin by deamination of arginine residues.";
RT      U. Immunol. 162:585-594(1999).
DR      EMBL, AF043380; AAC23559.1; -.
DR      GO: GO:0005198; F:structural molecule activity; IEA.
DR      InterPro: IPR003303; Filaggrin.
DR      Pfam: PF03516; Filaggrin; 2.
DR      PRINTS: PR00487; FILAGGRIN.
FT      NON_TER      1
FT      NON_TER      1
FT      NON_TER      1
SQ      SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF950 CRC64;

Query Match      79.5%; Score 70; DB 4; Length 322;
Best Local Similarity 77.8%; Pred. No. 0.00068;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      1  SHOESTYGRSGRSGSGS 18
      ||||| ||||| |||||
DB      305 SHOESTYGRSGRSGSGS 322

RESULT 10
ID      09V9Y3      PRELIMINARY;      PRT;      1284 AA.
AC      09V9Y3:
DT      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      CG11339-2A.
CN      CG11339.
OC      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBL_TaxID=7227;
      [1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.H., Plazek R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA      Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Bernson B.F., Bhandari D., Bolshakov S.,
RA      Bokoyva D., Borkan M.R., Bouck J., Brokstein P., Brotter P.,
RA      Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferris J., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegyan C.,
RA      Urali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrelet A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```


RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Viamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195 (2000).
 (2)
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Mathews B.B., Bayraktaroglu I., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Bertram B.P., Carlson J.W., Celniker S.E.,
 RA Clamp M.E., Dwyer R.A., Emmert D., Frise E., de Grey A.D.N.U.,
 RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnick F.,
 RA Whitefield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome,"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003777; AAF57145.2; -
 DR Flybase; FBgn0039841; CG11339.
 DR GO; GO:0005856; C:cytoskeleton; IEA.
 DR InterPro; IPR000289; Band 4.1.
 DR Pfam; PF00373; Band 4.1; 1.
 DR PRINTS; PRO0935; BAND41.
 DR SMART; SM00295; B41; 1.
 DR PROSITE; PS0057; FERM_3; 1.
 SQ SEQUENCE 1284 AA; 138966 MW; EDFA119838AA6447 CRC64;
 SO Query Match 58.0%; Score 51; DB 5; Length 1284;
 Best Local Similarity 61.1%; Pred. No. 5.8;
 Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 18
 Db 545 SDNESELRSRGSRSGS 562
 RESULT 11
 QP312 PRELIMINARY; PRT; 822 AA.
 ID QP312;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 D2 Related to nucleolar phosphoprotein.
 GN B12P1.10.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_Taxid=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL390091; CAB98213.1; -

DR PIR; T51049.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR005120; Smg-4_UPF3.
 DR Pfam; PF00076; trm; 1.
 DR Pfam; PF03467; Smg4_UPF3; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; 1.
 SQ SEQUENCE 822 AA; 86287 MW; E40A45DCC077245C CRC64;
 SO Query Match 55.7%; Score 49; DB 3; Length 822;
 Best Local Similarity 52.9%; Pred. No. 7.9;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGS 19
 Db 414 RESASGRTRGRGRGCT 430
 RESULT 12
 Q9UAY0 PRELIMINARY; PRT; 471 AA.
 ID Q9UAY0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE W03G1.5 protein.
 GN W03G1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson R., Baynes C., Barks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Crixton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkes T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Latching U., Lloyd C., McMurray A., Northmore B., O'Callaghan M.,
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Spratt J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*,"
 RL Nature 368:32-38 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A., Scheet P., Harper M.,
 RT "The sequence of *C. elegans* cosmid W03G1,"
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.,
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF125964; AD14753.1; -
 DR PIR; T33997; T33997.
 DR Wormpep; W03G1.5; CE17283.
 SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A955B CRC64;
 SO Query Match 54.5%; Score 48; DB 5; Length 471;
 Best Local Similarity 52.6%; Pred. No. 6.3;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 19
 Db 167 SSRSPSRGRGRGRGRSGS 185

RESULT 13

045386 PRELIMINARY; PRT; 306 AA.

AC 045386;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE F21H7.5 protein.

OS F21H7.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1] _SEQUENCE FROM N.A.

RA Gardner A.E.;

RN Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RP [2]

RX SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "genome sequence of the nematode C.elegans: A platform for investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL; Z93379; CAB07386.1; -.

DR PIR; T21220; T21220.

DR WormPep; F21H7.5; CE15877.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000535; MSP domain.

DR InterPro; IPR006962; Papd-like.

DR Pfam; PF006635; MSP domain; 1.

DR PROSITE; PS50202; MSP; 1.

SQ SEQUENCE 306 AA; 32421 MW; 830C17DB321D0C28 CRC64;

Query Match 53.4%; Score 47; DB 5; Length 306;

Best Local Similarity 44.4%; Pred. No. 5.8;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGRSG 18

DB 48 SHRSSXTNKRSGRSGRSG 65

RESULT 14

Q9NTA9 PRELIMINARY; PRT; 399 AA.

AC Q9NTA9;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein (Fragment).

OS DKFZP761E0323.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1] _SEQUENCE FROM N.A.

RA Bioecker H., Boecker M., Brandt P., Mewes H.W., Gassenhuber J., Wiemann S.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL137423; CAB70733.1; -.

DR PIR; T46259; T46259.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00271; Helicase_C; 1.

DR SMART; SM00490; HELIC_C; 1.

KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.

FT NON TER 1

SQ SEQUENCE 399 AA; 44075 MW; BC09B3A0E4CE5B96 CRC64;

Query Match 53.4%; Score 47; DB 4; Length 399;

Best Local Similarity 55.6%; Pred. No. 7.7;

Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGRSG 18

DB 336 SNRQRSGWSSGRSGRSG 353

RESULT 15

Q82C67 PRELIMINARY; PRT; 436 AA.

AC Q82C67;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative integral membrane protein.

GN SAV5487.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=33903;

RN [1] _SEQUENCE FROM N.A.

RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RT "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites."

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinoue M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.;

RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis."

RL Nat. Biotechnol. 21:526-531(2003).

DR EMBL; AP005043; BAC73199.1; -.

DR InterPro; IPR002194; Chaperonin_TCP-1.

DR PROSITE; PS00995; TCP1_3; 1.

KW Complete proteome.

SQ SEQUENCE 436 AA; 45359 MW; 35BD46B80C7282D0 CRC64;

Query Match 53.4%; Score 47; DB 16; Length 436;

Best Local Similarity 83.3%; Pred. No. 8.5;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GRSGRSGRSGS 19

DB 253 GRSGRSGRSGS 264

Search completed: September 28, 2004, 06:12:46

Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQSTXKRSXKRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of resules predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	AAW61515	AAW61515 Peptide c
2	84	95.5	19	AAW61507	AAW61507 Peptide c
3	84	95.5	19	AAW61517	AAW61517 Peptide c
4	84	95.5	19	AAW61511	AAW61511 Peptide c
5	84	95.5	19	AAW61505	AAW61505 Peptide c
6	84	95.5	19	AAW61514	AAW61514 Peptide c
7	84	95.5	19	AAW61516	AAW61516 Peptide c
8	84	95.5	19	AAW61516	AAW61516 Peptide c
9	84	95.5	19	AAW61516	AAW61516 Peptide c
10	84	95.5	19	AAW61516	AAW61516 Peptide c
11	84	95.5	19	AAW61516	AAW61516 Peptide c
12	84	95.5	19	AAW61516	AAW61516 Peptide c
13	84	95.5	19	AAW61516	AAW61516 Peptide c
14	84	95.5	19	AAW61516	AAW61516 Peptide c
15	84	95.5	19	AAW61516	AAW61516 Peptide c
16	84	95.5	19	AAW61516	AAW61516 Peptide c
17	84	95.5	19	AAW61516	AAW61516 Peptide c
18	84	95.5	19	AAW61516	AAW61516 Peptide c
19	84	95.5	19	AAW61516	AAW61516 Peptide c
20	84	95.5	19	AAW61516	AAW61516 Peptide c
21	84	95.5	19	AAW61516	AAW61516 Peptide c
22	84	95.5	19	AAW61516	AAW61516 Peptide c
23	84	95.5	19	AAW61516	AAW61516 Peptide c
24	84	95.5	19	AAW61516	AAW61516 Peptide c
25	84	95.5	19	AAW61516	AAW61516 Peptide c

26	47	53.4	574	4	AAW41855	AAW41855 Human pol
27	47	53.4	752	4	AAW23535	AAW23535 Novel hum
28	44	50.0	106	3	AAW33285	AAW33285 Zea mays
29	44	50.0	108	3	AAW33268	AAW33268 Zea mays
30	44	50.0	138	4	AAW50949	AAW50949 Propionib
31	44	50.0	138	6	AAW47468	AAW47468 Propionib
32	44	50.0	233	3	AAW33195	AAW33195 Zea mays
33	43.5	49.4	286	2	AAW85441	AAW85441 Neurotrop
34	43.5	49.4	19938	6	AAW76679	AAW76679 Streptomy
35	43	48.9	92	4	AAW60843	AAW60843 Propionib
36	43	48.9	92	6	AAW57362	AAW57362 Propionib
37	43	48.9	477	4	AAW608221	AAW608221 Novel hum
38	43	48.9	641	4	AAW39110	AAW39110 Novel hum
39	43	48.9	649	6	AAW48672	AAW48672 Protein e
40	43	48.9	694	2	AAW1267	AAW1267 Drosophila
41	43	48.9	694	4	AAW71797	AAW71797 Drosophila
42	43	48.9	982	2	AAW13320	AAW13320 Murine Na
43	43	48.9	982	6	AAW05134	AAW05134 Human exp
44	43	48.9	1482	7	AAW58671	AAW58671 Rat Prote
45	43	48.9	1482	7	AAW58675	AAW58675 Rat Prote

ALIGNMENTS

RESULT 1
AAW61515
ID AAW61515 standard; peptide, 19 AA.
AC AAW61515;
AA AAW61515;
DT 26-OCT-1998 (first entry)
DX Peptide cfa, based on cDNA of a profilaggrin repeat.
DS Peptide cfa, based on cDNA of a profilaggrin repeat.
DX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
XX WC9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WC-ND000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6, 19pp; English.
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX Sequence 19 AA;

Query Match 96.6%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
DB 1 SHOESTXGRSGRSGRSGS 19

RESULT 2

AAW61507 ID AAW61507 standard; peptide; 19 AA.
XX AAW61507;
AC AAW61507;
XX 26-OCT-1998 (first entry)
DT Peptide cfc3, based on cDNA of a profilaggrin repeat.
XX
DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Modified-site 11 /note="Citruiline"
FT
XX
PN W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97MO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
DR
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
CC
XX Sequence 19 AA;
SQ

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 1.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
DB 1 SHOESTXGRSGRSGRSGS 19

RESULT 3

AAW61517 ID AAW61517 standard; peptide; 19 AA.
XX AAW61517;
AC AAW61517;
XX 26-OCT-1998 (first entry)
DT Peptide cfcQ, based on cDNA of a profilaggrin repeat.
XX
DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97MO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
DR
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
CC
XX Sequence 19 AA;
SQ

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
DB 1 SHOESTXGRSGRSGRSGS 19

RESULT 4

AAW61511 ID AAW61511 standard; peptide; 19 AA.
XX AAW61511;
AC AAW61511;
XX 26-OCT-1998 (first entry)
DT Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citnulline"
FT Modified-site 11 /note= "Citnulline"
FT Modified-site 11 /note= "Citnulline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-06;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SHQESTXGRSGRSGS 19
XX 1 SHQESTXGRSGRSGS 19
XX
XX Db 1 SHQESTXGRSGRSGS 19
XX
XX RESULT 5
XX AAW61505
XX ID AAW61505 standard; peptide; 19 AA.
XX
XX AAW61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citnulline"
FT Modified-site 7 /note= "Citnulline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX

```

```

PF 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 1.2e-06;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 SHQESTXGRSGRSGS 19
XX 1 SHQESTXGRSGRSGS 19
XX
XX Db 1 SHQESTXGRSGRSGS 19
XX
XX RESULT 6
XX AAW61514
XX ID AAW61514 standard; peptide; 19 AA.
XX
XX AAW61514;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cf, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX

```

PS Disclosure; Page 6; 19pp; English.

XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.2e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 7
 ID AAM61516 standard; peptide; 19 AA.
 XX AC AAM61516;
 XX DT 26-OCT-1998 (first entry)
 XX DE Peptide cFE, based on cDNA of a profilaggrin repeat.
 XX KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX KW monoclonal antibody.
 XX OS Synthetic.
 XX OS Homo sapiens.
 XX PN WO922503-A2.
 XX PD 28-MAY-1998.
 XX PF 14-NOV-1997; 97WO-NL000624.
 XX PR 15-NOV-1996; 96NL-01004539.
 XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX PI Van Venrooij WTM, Schellekens GA, Raats JWH, Host RMA;
 DR WPI; 1998-398613/34.
 XX PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX PS Disclosure; Page 6; 19pp; English.
 XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 8
 ID ABB97605 standard; protein; 1467 AA.
 XX AC ABB97605;
 XX DT 27-JUN-2002 (first entry)
 XX DE Novel human protein SEQ ID NO: 873.
 XX KM Human; anti-naemic; vulnerary; anti-inflammatory; immunomodulator;
 XX KM antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 XX KM neuroprotective; antiparkinsonian; protein therapy; EST;
 XX KM expressed sequence tag.
 XX OS Homo sapiens.
 XX PN WO200222660-A2.
 XX PD 21-MAR-2002.
 XX PF 10-SEP-2001; 2001WO-US026015.
 XX PR 11-SEP-2000; 2000US-00659671.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Zhou P, Aseudi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX DR WPI; 2002-292408/33.
 XX DR N-PSDB; ABB97605.
 XX PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX PS Example 2; SEQ ID NO 873; 503pp; English.
 XX CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate the immune system e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

XX SQ Sequence 1467 AA;

Query Match 95.5%; Score 84; DB 5; Length 1467;
 Best Local Similarity 89.5%; Pred. No. 7.3e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 Db 1097 SHOESTRGRSGRSGSGS 1115

RESULT 9
 ID AAM61506 standard; peptide; 19 AA.
 XX AC AAM61506;

DT 26-OCT-1998 (first entry)
 XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 DB 1 SHQESTRGRSGRSGSGS 19
 RESULT 10
 AAW61508
 ID AAW61508 standard; peptide; 19 AA.
 XX
 XX AAW61508;
 AC
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc4, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 13 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 DB 1 SHQESTRGRSGRSGSGS 19
 RESULT 11
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 XX AAW61512;
 AC
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX

XX	PR	15-NOV-1996;	96NL-01004539.
XX	PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
XX	PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
XX	PI	Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;	
XX	DR	WPI; 1998-398613/34.	
XX	PT	Peptide derived from an antigen recognised by autoantibodies - is	
XX	PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
XX	PS	used in diagnosis of the disease.	
XX	PS	Disclosure; Page 6; 19pp; English.	
CC	CC	Sequences AAM61505-W61520 are peptides derived from the C-terminal end of	
CC	CC	the profilaggrin antigen which is recognised by autoantibodies from	
CC	CC	patients with rheumatoid arthritis (RA). This peptide is reactive with a	
CC	CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	CC	The peptides were created by using standard solid phase synthesis, which	
CC	CC	produced them as peptide amides. These sequences may be used in the	
CC	CC	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as	
CC	CC	for obtaining polyclonal and monoclonal antibodies	
SQ	SQ	Sequence 19 AA;	
QY	QY	Query Match 88.6%; Score 78; DB 2; Length 19;	
Dd	Dd	Best Local Similarity 89.5%; Pred.No.1.le-05;	
		Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0	
		1 SHOESTXGRSGRSGRS 15	
		1 SHOESTXGRSGRSGRS 19	
RESULT 12			
ID	AA61509	AA61509 standard, peptide, 19 AA.	
AC	AA61509;		
DT	26-OCT-1998	(first entry)	
DE	Peptide cfc5, based on cDNA of a profilaggrin repeat.		
KM	Antigen, autoantibody; rheumatoid arthritis; RA; autoimmune antibody;		
KM	solid phase synthesis; peptide amide; polyclonal antibody;		
KM	monoclonal antibody.		
OS	Synthetic.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	Modified-site	16	
FT	/note="Citrulline"		
PN	MO9822503-A2.		
PD	28-MAY-1998.		
PF	14-NOV-1997;	97WO-NL000624.	
PR	15-NOV-1996;	96NL-01004539.	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.		
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.		
PI	Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;		
DR	WPI; 1998-398613/34.		
PT	Peptide derived from an antigen recognised by autoantibodies - is		

[illegible]

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 1 SHOESTXGRSGRSGSGS 19
 DB 1 SHOESTXGRSGRSGSGS 19

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 DT 26-OCT-1998 (first entry)
 DE Peptide cfc6, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

MO9822503-A2;
 PD 28-MAY-1998.

PF 14-NOV-1997; 97MO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHOESTXGRSGRSGSGS 19
 1 SHOESTXGRSGRSGSGS 19
 DB 1 SHOESTXGRSGRSGSGS 19

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 DT 26-OCT-1998 (first entry)
 DE Peptide XI based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

MO9822503-A2.
 PD 28-MAY-1998.

PF 14-NOV-1997; 97MO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 3.6e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGSGS 19
 4 HOESTXGRSGRSGSGS 21
 DB 4 HOESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:18
 Job time : 56.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)

54.782 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-3
4	84	95.5	19	9	US-09-308-150-7
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	78	88.6	19	9	US-09-308-150-2
9	78	88.6	19	9	US-09-308-150-4
10	78	88.6	19	9	US-09-308-150-5
11	78	88.6	19	9	US-09-308-150-6
12	78	88.6	19	9	US-09-308-150-8
13	78	88.6	19	9	US-09-308-150-9
14	75	85.2	21	9	US-09-308-150-10
15	75	85.2	22	9	US-09-747-029A-22

16	51	58.0	477	15	US-10-161-927-62	Sequence 62, Appl
17	49	55.7	179	12	US-10-425-114-53545	Sequence 53545, A
18	48	54.5	123	16	US-10-767-701-32436	Sequence 32436, A
19	47	53.4	164	16	US-10-437-963-112419	Sequence 112419, A
20	46	53.4	436	14	US-10-156-761-103022	Sequence 103022, A
21	45	52.3	243	16	US-10-437-963-103033	Sequence 103033, A
22	45	51.1	612	12	US-10-424-599-230819	Sequence 230819, A
23	44	50.0	133	16	US-10-437-963-136367	Sequence 136367, A
24	44	50.0	155	16	US-10-437-963-161956	Sequence 161956, A
25	44	50.0	653	16	US-10-437-963-198539	Sequence 198539, A
26	44	50.0	653	9	US-09-746-801A-13	Sequence 13, Appl
27	44	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
28	44	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
29	44	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
30	43.5	49.4	19723	15	US-10-084-846A-5	Sequence 5, Appl1
31	43	48.9	232	16	US-10-437-963-183394	Sequence 183394, A
32	43	48.9	536	16	US-10-437-963-149523	Sequence 149523, A
33	43	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
34	43	48.9	649	12	US-10-282-122A-76596	Sequence 76596, A
35	43	48.9	694	14	US-10-152-548-2	Sequence 2, Appl1
36	42	47.7	436	15	US-10-369-493-8178	Sequence 8178, Ap
37	42	47.7	506	16	US-10-437-963-185974	Sequence 185974, A
38	42	47.7	588	16	US-10-437-963-177280	Sequence 177280, A
39	41	46.6	24	9	US-09-971-543-12	Sequence 12, Appl
40	41	46.6	50	14	US-10-029-386-27692	Sequence 27692, A
41	41	46.6	103	12	US-10-424-599-257534	Sequence 257534, A
42	41	46.6	136	16	US-10-437-963-178685	Sequence 178685, A
43	41	46.6	178	12	US-10-425-114-69976	Sequence 69976, A
44	41	46.6	184	9	US-09-925-300-1619	Sequence 1619, Ap
45	41	46.6	184	12	US-10-296-115-763	Sequence 763, App

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00524
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-09-308-150-12
OTHER INFORMATION: Known CDNA sequences of human profilaggrin
Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0;
Gy 1 SHQESTXGRSGRSGSGS 19

Db 1 SHOESTXGRSGRSGSGS 19

RESULT 2
US-09-308-150-1
Sequence 1, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL.1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 6.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 3
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL.1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 6.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 4
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL.1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 5
US-09-308-150-11
Sequence 11, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL.1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 11
LENGTH: 19

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1996-11-15
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-11

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Oy      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTGRGRSGRSGSGS 19

```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-13

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Oy      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTGRGRSGRSGSGS 19

```

```

RESULT 7
US-09-308-150-14

```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-14

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Oy      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTGRGRSGRSGSGS 19

```

```

RESULT 8
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-2

```

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGSGS 19
|||
Db 1 SHQESTXGRSGRSGSGS 19

RESULT 9
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4

LENGTH: 19
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGSGS 19
|||
Db 1 SHQESTXGRSGRSGSGS 19

RESULT 10
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5

LENGTH: 19
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGSGS 19
|||
Db 1 SHQESTXGRSGRSGSGS 19

RESULT 11
US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6

LENGTH: 19
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGSGS 19
|||
Db 1 SHQESTXGRSGRSGSGS 19

RESULT 12
US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. NO. 5,1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

```

```

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. NO. 5,1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. NO. 0.00016;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 HOESTXGRSGRSGSGS 19
DB 4 HOESTXGRSGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NPU500 INNS.031
; CURRENT APPLICATION NUMBER: US/09/747,029A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
; NAME/KEY: MOD_RES
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22
    
```

```

Query Match      85.2%; Score 75; DB 9; Length 22;
Best local similarity 88.9%; Pred. No. 0.00017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
    
```

```

Qy      2 HQESTXGRSXGRSGRSGS 19
          |||||
Db      5 HQESTXGRSRGRGCRSGS 22
    
```

```

Search completed: September 28, 2004, 07:28:51
Job time : 111.526 secs
    
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQSTYKGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/FCRUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	45	51.1	518	US-09-252-991A-25967
2	44	50.0	395	US-09-252-991A-21702
3	44	50.0	483	US-09-252-991A-19015
4	44	50.0	653	US-09-513-057C-13
5	43	48.9	629	US-09-252-991A-22901
6	42	47.7	85	US-09-252-991A-29339
7	41.5	47.2	96	US-09-252-991A-19216
8	41	46.6	120	US-09-252-991A-17025
9	41	46.6	237	US-09-252-991A-17927
10	41	46.6	255	US-09-252-991A-19598
11	41	46.6	256	US-09-252-991A-26244
12	41	46.6	406	US-09-252-991A-19857
13	41	46.6	421	US-09-252-991A-33326
14	41	46.6	722	US-09-252-991A-33326
15	41	46.6	724	US-09-252-991A-33326
16	41	46.6	745	US-09-252-991A-33326
17	41	46.6	2237	US-09-252-991A-33326
18	40	45.5	123	US-09-252-991A-28146
19	40	45.5	213	US-09-252-991A-32093
20	40	45.5	261	US-09-252-991A-29394
21	40	45.5	429	US-09-252-991A-29394
22	40	45.5	429	US-10-109-854-5
23	40	45.5	429	US-09-252-991A-19026
24	40	45.5	468	US-09-252-991A-19026
25	40	45.5	484	US-09-252-991A-19026
26	40	45.5	508	US-09-252-991A-19026
27	40	45.5	722	US-08-817-832B-32

28	39.5	44.9	191	4	US-09-252-991A-18626	Sequence 18626, A
29	39	44.3	147	4	US-09-252-991A-22232	Sequence 22232, A
30	39	44.3	148	4	US-09-252-991A-25505	Sequence 25505, A
31	39	44.3	176	4	US-09-252-991A-29512	Sequence 29512, A
32	39	44.3	195	4	US-09-252-991A-30082	Sequence 30082, A
33	39	44.3	409	4	US-09-252-991A-14217	Sequence 14217, A
34	39	44.3	491	4	US-09-252-991A-7836	Sequence 7836, A
35	39	44.3	778	4	US-09-252-991A-18605	Sequence 18605, A
36	39	44.3	1293	4	US-09-170-496D-252	Sequence 252, A
37	39	44.3	1293	4	US-09-364-425B-57	Sequence 2204, A
38	38.5	43.8	257	4	US-09-252-991A-22204	Sequence 30849, A
39	38.5	43.8	388	4	US-09-252-991A-30849	Sequence 22, A
40	38	43.2	26	4	US-09-297-981-22	Sequence 20, A
41	38	43.2	46	4	US-09-297-981-20	Sequence 10, A
42	38	43.2	223	1	US-07-667-276A-10	Sequence 32606, A
43	38	43.2	279	4	US-09-252-991A-32606	Sequence 21789, A
44	38	43.2	299	4	US-09-252-991A-21789	Sequence 20245, A
45	38	43.2	369	4	US-09-252-991A-20245	

ALIGNMENTS

RESULT 1
US-09-252-991A-25967
Sequence 25967, Application US/09252991A
Parent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25967
LENGTH: 518
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25967

Query Match 51.1% Score 45; DB 4; Length 518;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 STGRSGRSGRSG 18
DB 412 SGGGRSGRSGRSG 425
RESULT 2
US-09-252-991A-21702
Sequence 21702, Application US/09252991A
Parent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21702
LENGTH: 395
TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21702
Query Match          50.0%; Score 44; DB 4; Length 395;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTYGRXGRSGRSG 18
DB 251 SHQASERGRDHSRGRPG 268

RESULT 3
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18015
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015
Query Match          50.0%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTYGRXGRSGRSG 18
DB 94 HAEQDPGAGTGRAGRAG 110

RESULT 4
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
US-09-513-057C-13
Query Match          50.0%; Score 44; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTYGRXGRSGRSG 19
DB 573 SRQVSTASSASGREGISGS 591

RESULT 5
US-09-252-991A-22901
; Sequence 22901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22901
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22901
Query Match          48.9%; Score 43; DB 4; Length 629;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTYGRXGRSGRSGS 19
DB 272 RQSHRGRAGRAGRGRGS 288

RESULT 6
US-09-252-991A-29339
; Sequence 29339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29339
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29339
Query Match          47.7%; Score 42; DB 4; Length 85;
Best Local Similarity 80.0%; Pred. No. 6;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GRXGRSGRSGS 17
DB 73 GRAGRSGRSGS 82

RESULT 7
US-09-252-991A-19216
; Sequence 19216, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19216
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19216
Query Match          48.9%; Score 43; DB 4; Length 629;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTYGRXGRSGRSGS 19
DB 272 RQSHRGRAGRAGRGRGS 288

RESULT 6
US-09-252-991A-29339
; Sequence 29339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29339
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29339
Query Match          47.7%; Score 42; DB 4; Length 85;
Best Local Similarity 80.0%; Pred. No. 6;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GRXGRSGRSGS 17
DB 73 GRAGRSGRSGS 82

RESULT 7
US-09-252-991A-19216
; Sequence 19216, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19216
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19216
Query Match          48.9%; Score 43; DB 4; Length 629;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTYGRXGRSGRSGS 19
DB 272 RQSHRGRAGRAGRGRGS 288
```

PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19216
LENGTH: 96
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19216

Query Match 47.2%; Score 41.5; DB 4; Length 96;
Best Local Similarity 62.5%; Pred. No. 8.2;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 5 STXGR-SXGRSGSGS 19
DB 75 SATGRGAGGAGRSGS 90

RESULT 8
US-09-252-991A-17025
Sequence 17025, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17025
LENGTH: 120
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17025

Query Match 46.6%; Score 41; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 SXGRSGSGS 19
DB 7 SAGRAGRSGS 16

RESULT 9
US-09-252-991A-17927
Sequence 17927, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17927
LENGTH: 237
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (237)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-17927

Query Match 46.6%; Score 41; DB 4; Length 237;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GRGXGRSGRS 17
DB 24 GRGGRGREGRA 33

RESULT 10
US-09-252-991A-19598
Sequence 19598, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19598
LENGTH: 255
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19598

Query Match 46.6%; Score 41; DB 4; Length 255;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 GRGXGRSGRS 19
DB 64 GRDGRGAGRAGA 75

RESULT 11
US-09-252-991A-26244
Sequence 26244, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26244
LENGTH: 256
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26244

Query Match 46.6%; Score 41; DB 4; Length 256;
Best Local Similarity 47.1%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSG 18
DB 97 HRRRTAGRAGRRPRRG 113

RESULT 12

```

US-09-252-991A-19857
; Sequence 19857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19857
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19857

Query Match          46.6%; Score 41; DB 4; Length 406;
Best Local Similarity 44.4%; Pred. No. 43;
Matches      8; Conservative 2; Mismatches      8; Indels 0; Gaps 0;

QY      2 HSESTXGRSXYGRSGRSGS 19
Db      127 HGSRRGRDRAGRGGRGCA 144

RESULT 13
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326

Query Match          46.6%; Score 41; DB 4; Length 421;
Best Local Similarity 47.1%; Pred. No. 44;
Matches      8; Conservative 1; Mismatches      8; Indels 0; Gaps 0;

QY      2 HSESTXGRSXYGRSGRSGS 18
Db      13 HARTGTGTGRRRRRSG 29

RESULT 14
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31

```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match          46.6%; Score 41; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 77;
Matches      9; Conservative 0; Mismatches      6; Indels 0; Gaps 0;

QY      5 STXGRSXYGRSGRSGS 19
Db      562 SPSGHSQGRGRGASGS 576

RESULT 15
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-890-2

Query Match          46.6%; Score 41; DB 4; Length 724;
Best Local Similarity 60.0%; Pred. No. 77;
Matches      9; Conservative 0; Mismatches      6; Indels 0; Gaps 0;

QY      5 STXGRSXYGRSGRSGS 19
Db      564 SPSGHSQGRGRGASGS 578

```

Search completed: September 28, 2004, 06:26:45
Job time : 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88

Sequence: 1 SHOESTXGRSGXSGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	2248	2 A35938	profilaggrin - hum
2	78	88.6	416	2 A32947	profilaggrin precursor
3	69	78.4	591	2 A45135	profilaggrin - hum
4	52	59.1	506	1 W2W47	E2 protein - human
5	49	55.7	822	1 T51049	related to nucleol
6	46.5	52.8	135	2 A46598	RNA-binding protei
7	45	51.1	2420	2 A84652	hypothetical prote
8	44	50.0	638	2 I53169	cytochrome 2 - hu
9	44	50.0	800	2 T02852	probable membrane
10	44	50.0	1829	2 T35681	probable sensory h
11	43	48.9	306	2 T21220	hypothetical prote
12	43	48.9	694	2 S71786	wingless receptor
13	43	48.9	849	2 A96592	hypothetical prote
14	43	48.9	1015	2 T42013	frequency clock pr
15	43	48.9	1507	2 B47328	natural killer cel
16	42	47.7	136	2 T35632	probable transpos
17	42	47.7	176	2 F71370	probable single-st
18	42	47.7	772	2 T27512	hypothetical prote
19	41	46.6	471	2 T33997	hypothetical prote
20	41	46.6	825	2 UC4163	DNA-binding protei
21	41	46.6	836	2 G84727	probable DNA topoi
22	41	46.6	843	1 JDUVLR	DNA-directed DNA p
23	40.5	46.0	311	2 T45683	hypothetical prote
24	40	45.5	128	2 A72654	glycine-rich prote
25	40	45.5	129	2 S14984	probable ribosoma
26	40	45.5	150	2 B71185	lsu ribosomal prot
27	40	45.5	151	2 F75145	hypothetical prote
28	40	45.5	163	2 T23076	hypothetical prote
29	40	45.5	269	2 B38095	T-cell-specific tr

30	40	45.5	272	2 A38900	T-cell-specific tr
31	40	45.5	303	2 JH0401	T-cell-specific tr
32	40	45.5	312	2 A31846	130K paracrystall
33	40	45.5	373	2 T02976	probable DNA bindi
34	40	45.5	399	2 T46259	hypothetical prote
35	40	45.5	439	2 A38160	chd protein - Esc
36	40	45.5	462	4 S33798	FUS/CHOP mutant fu
37	40	45.5	526	1 S33799	RNA-binding protei
38	40	45.5	827	2 T13468	DNA-directed DNA p
39	40	45.5	842	1 JDUVRS	DNA-directed DNA p
40	40	45.5	854	2 S61164	repair/recombinati
41	40	45.5	1232	2 S40766	hypothetical prote
42	40	45.5	2176	2 T13806	coucan gene protei
43	39	44.3	110	2 B70601	hypothetical prote
44	39	44.3	259	2 C69122	hypothetical prote
45	39	44.3	313	2 A28444	flaggrin precursor

ALIGNMENTS

RESULT 1

A35938

profilaggrin - human (fragments)

C:Species: Homo sapiens (man)

C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

C:Accession: A35938

R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.

Biochemistry 29, 9432-9440, 1990

A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.

A:Reference number: A35938; MUID:91064347; PMID:2248957

A:Accession: A35938

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2248 <GN>

A:Cross-references: GB:U02929

A:Genetics:

A:Gene: GDB:FLG

A:Cross-references: GDB:119912; OMIM:135940

A:Map position: 1q21-1q21

C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

F:246-566/Region: flaggrin repeat

F:570-893/Region: flaggrin repeat

F:1074-1397/Region: flaggrin repeat

F:1573-1896/Region: flaggrin repeat

Query Match 95.5%; Score 84; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19

DB 227 SHOESTGRSGXSGRSGS 245

RESULT 2

A32947

flaggrin precursor - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999

C:Accession: A32947

R:McKinnley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc

Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989

A:Title: Characterization of a cDNA clone encoding human flaggrin and localization of cl

A:Reference number: A32947; MUID:89296901; PMID:2740331

A:Accession: A32947

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <NC>

A:Cross-references: GB:M4355; NID:9182604; PID:AA52454.1; PID:9182605

A:Note: the authors translated the codon CAC for residue 186 as Gln, and AAT for residue

C:Genetics:

A;Gene: GDB:FIG
A;Cross-references: GDB:119912; OMIM:135940
A;Map position: 1q21-1q21
C;Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C;Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match	88.6%;	Score 78;	DB 2;	Length 416;
Best Local Similarity	84.2%;	Pred. No. 3.8e-05;		
Matches 16;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

QY 1 SHQESTXGRSRGXSGRSGS 19
Db 7 SHQESTRGRSAGRSGRSGS 25

RESULT 3

Profilaggrin human (fragment)
C/Species: Homo sapiens (man)
C/Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1995
C/Accession: A45135
R/Prepared: R.B.; Haydock, P.V.; Fleckman, P.; Nirunskis, I.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992
A/Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A/Reference number: A45135; MUID:93054736; PMID:1428717
A/Accession: A45135
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-591 <PFE>
A/Cross-references: GB:L01089; GB:M90967; NID:G190408; PIDN:AAA60177.1; PID:G553621
A/Note: sequence extracted from NCBI backbone (NCBI:P:118773)
C/Genetics:
C/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMTM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
C/49-81/Domain: calmodulin repeat homology <EF2>

Query Match	78.4%;	Score 69;	DB 2;	Length 591;
-------------	--------	-----------	-------	-------------

Matches	15;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY	1	SHQESTXGRSRGXSGRGS	19
Db	449	SHQESTRGRSGERSGRGS	46

RESULT 4

E2 protein - human papillomavirus type 47
C/Species: human papillomavirus type 47
A/Note: host Homo sapiens (man)
C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C/Accession: D35324
R/Kiyono, T.; Adachi, A.; Iehibashi, M.
Virology 177, 401-405, 1990
A/Title: Genome organization and taxonomic position of human papillomavirus type 47 inferred from DNA sequencing
A/Reference number: A35324; MUID:90284511; PMID:212112
A/Accession: D35324
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-506 <K1Y>
A/Cross-references: GB:U32305; NID:g333062; PID:AAA6979.1; PID:g333067
C/Superfamily: papillomavirus E2 protein
C/Keywords: DNA binding; early protein; transcription regulation

```

Query Match      59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.77;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0

```

Db 342 SREGNTRGGRGRQGRAGS 360

RESULT 5

related to nucleolar phosphoprotein [imported] - *Neurospora crassa*

C; Species: *Neurospora crassa*

C;Accession: T51049

submitted to the Protein Sequence Database, July 2000
 A: Reference number: Z25286

A;Accession: T51049

A;Molecule type: DNA
A;Residues: 1-822 <SC

A:Experimental source: BAC clone B12F1; strain OR74A
A;Cross-references: EMBL:AL390091; GSDB:GN00116; NCSP:B12F1.10

A:Gene: NCSP:B12F1.10

A; Introns: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;

```

      EDC: Local similarity    92.08;   Area: NO: 0.0;
Matches    9; Conservative    3; Mismatches    5; Indels    0; Gaps    0;

```

```
QY      3 QESTXGRSRGXSGRSGS 19
      :|||:|||||:
Db      414 RESASGRTGRGRGGT 430
```

RESULT 6

RNA-binding protein 1 - fruit fly (*Drosophila melanogaster*)
C/Species: *Drosophila melanogaster*
C/Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 24-Sep-1999
C/Accession: A46398
R/Kim, Y.J.; Zuo P.; Manley, J.L.; Baker, B.S.
Genes Dev. 6, 2869-2879, 1992
A/Title: The *Drosophila* RNA-binding protein RBP1 is localized to transcriptionally active
A/Reference number: A46398; MWID:94040720; PMID:1340470
A/Accession: A46398
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-135 <KIM>
A/Cross-references: GB:I04929; NID:g158223; PIDN:AAA28950.1; PID:g158224
A/Note: authors translated the codon GCT for residue 65 as Arg
C/Genetics: rbp1
A/Gene: FlyBase:rbp1
A/Cross-references: FlyBase:FBgn0010252
A/Intons: 106/2
C/Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C/12-74/Domain: ribonucleoprotein repeat homology <RNM3>

Query Match	52.88; Score 46.5; DB 2; Length 135;
-------------	--------------------------------------

Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

4 ESTXGRSR-----GXSGRSGS 19

Db 80 EMSGRSDRRGEGSSGRSGS 102

RESULT 7

hypothetical protein At2g25730 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001
C/Accession: A84652
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Nishizawa, T.; Holt, T.A.; Bowers, J.E.; Bowman, L.P.; Brice, M.C.; Brown, G.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.; Kim, H.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanken, S.E.; Umayam, L.; Tallon, L.

enes, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A/Reference number: A84420; MID:20083487; PMID:10617197
 A/Accession: A84652
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-2420 <STO>
 A/Cross-references: GB:AF002093; NID:94874311; PIDN:AA031373.1; GSPDB:GN00139
 C/Genetics:
 A/Map position: 2

Query Match 51.1%; Score 45; DB 2; Length 2420;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSGS 19
 |||||
 Db 1353 STSGRSGXGCGS 1367

RESULT 8
 153169
 CytoKeratin 2 - human
 C/Species: Homo sapiens (man)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
 C/Accession: 153169
 R/Collin, C.; Ouhayoun, J.
 A/Title: Suprabasal marker proteins distinguishing keratinizing squamous epithelia: Cyto
 A/Reference number: 153169; MID:93114504; PMID:1282112
 A/Accession: 153169
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-638 <RBS>
 A/Cross-references: GB:M99063; NID:9181389; PIDN:AAA5746.1; PID:9181390
 C/Superfamily: cytoskeletal keratin

Query Match 50.0%; Score 44; DB 2; Length 638;
 Best Local Similarity 60.0%; Pred. No. 19;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTGRSGXSGRSGS 17
 |||||
 Db 7 KKSFSGRSGSGS 21

RESULT 9
 T02852
 Probable membrane protein L1439.4 (imported) - Leishmania major (strain Friedlin)
 C/Species: Leishmania major
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C/Accession: T02852; H81462
 R/Wyler, P.J.
 A/Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
 submitted to the EMBL Data Library, May 1998
 A/Reference number: Z14740
 A/Accession: T02852
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-800 <MTL>
 A/Cross-references: EMBL:AF001274; NID:93264850; PID:92266920
 R/Wyler, P.J.; Andelman, L.; deVos, T.; Hixson, G.; Kiser, F.; Lemley, C.; Magness, C.;
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A/Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A/Reference number: A81453; MID:99178987; PMID:10077609
 A/Accession: H81462
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-800 <PVL>
 A/Cross-references: GB:AE001274; NID:93264850; PIDN:AA024675.1; PID:92266920; GSPDB:GN00
 A/Experimental source: strain WHOM/IL/81/Friedlin

C/Genetics:
 A/Gene: L1439.4
 A/Map position: 1

Query Match 50.0%; Score 44; DB 2; Length 800;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQESTGRSGXSGRSGS 19
 |||||
 Db 429 HRDGVRLSTTAAGSGS 446

RESULT 10
 T35681
 Probable sensory histidine kinase - Streptomyces coelicolor
 C/Species: Streptomyces coelicolor
 C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C/Accession: T35681
 R/Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1998
 A/Reference number: Z21587
 A/Accession: T35681
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1829 <HAR>
 A/Cross-references: EMBL:AL031031; PIDN:CA019849.1; GSPDB:GN00070; SCOEDB:SC7C7.03
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SCOEDB:SC7C7.03

Query Match 50.0%; Score 44; DB 2; Length 1829;
 Best Local Similarity 60.0%; Pred. No. 53;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSGS 19
 |||||
 Db 21 SNQGRSGXSGRSGS 35

RESULT 11
 T21220
 Hypothetical protein F21H7.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C/Accession: T21220
 R/Gardner, A.
 submitted to the EMBL Data Library, March 1997
 A/Reference number: Z19393
 A/Accession: T21220
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-306 <WII>
 A/Cross-references: EMBL:Z93379; PIDN:CA07568.1; GSPDB:GN00023; CBSP:F21H7.5
 A/Experimental source: clone F21H7
 C/Genetics:
 A/Gene: CBSP:F21H7.5
 A/Map position: 5
 A/Insertions: 28/3; 146/3; 232/3
 C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTGRSGXSGRSGS 18
 |||||
 Db 48 SHRSKTKRSGSGSGS 65

RESULT 12
 S71786
 wingless receptor precursor dfz2 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
 C>Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 C/Accession: S71786; S78444
 R/Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang, Y.; Macke, J.P.; Andrew, Nature 382, 223-230, 1996
 A>Title: A new member of the frizzled family from Drosophila functions as a wingless receptor
 A/Reference number: S71786; MUID:96353971; PMID:8717036
 A/Accession: S71786
 A>Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-694 <BHA>
 A/Cross-references: EMBL:U65589
 A/Note: mRNA was also sequenced
 R/Bhanot, P.; Wang, Y.; Nathans, J. submitted to the EMBL Data Library, July 1996
 A/Reference number: S78444
 A/Accession: S78444
 A/Molecule type: DNA
 A/Residues: 1-416, 'T', 418-694 <BHW>
 A/Cross-references: EMBL:U65589; NID:91518050; PIDN:AA047273.1; PID:91518051
 C/Genetics:
 A:Gene: dfz2
 A/Cross-references: FlyBase:FBgn0016797
 C/Superfamily: fruit fly frizzled protein
 C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 30;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHQSTXGRSGRSGSGS 19
 DB 182 SYTAGSGSGSGSGSGS 200

RESULT 13
 A96592
 hypochloral protein P14Q21.55 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: A96592
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chitt, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anssen, N.F.; Hughes, B.; Hutzar, L. Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: A96592
 A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-849 <STO>
 A/Cross-references: GB:AE005173; NID:911055759; PIDN:AA028231.1; GSPDB:GNC00141
 C/Genetics:
 A:Gene: F14Q21.55
 A/Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 5 STXGRSGRSGRSGS 19
 DB 736 SSANRSGRSGRSGS 750

RESULT 14
 T42013
 frequency clock protein - Creopus spinulosus

C:Species: Creopus spinulosus
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
 C/Accession: T42013
 R/Lewis, M.T.; Morgan, L.W.; Feldman, J.F. Mol. Gen. Genet. 253, 401-414, 1997
 A>Title: Analysis of frequency (frq) clock protein homologs: evidence for a helix-turn-helix
 A/Reference number: 222024; MUID:97188515; PMID:9037100
 A/Accession: T42013
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1015 <LEW>
 A/Cross-references: EMBL:U25850; NID:9852501; PID:9852502; PIDN:AAA68072.1
 C/Genetics:
 A:Gene: frq
 A/Introns: 100/1

Query Match 48.9%; Score 43; DB 2; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 44;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 SHQSTXGRSGRSGRSGS 19
 DB 583 NHRQKTGHSTGDSGSGN 601

RESULT 15
 B47328
 natural killer cell tumor-recognition protein - mouse
 N/Alternate names: cyclophilin-related NK-tumor recognition protein; natural killer-tumor
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 05-Nov-1999
 C/Accession: B47328; I77662
 R/Anderson, S.K.; Gallinger, S.; Roder, J.; Frey, J.; Young, H.A.; Ortaldo, J.R. Proc. Natl. Acad. Sci. U.S.A. 90, 542-546, 1993
 A>Title: A cyclophilin-related protein involved in the function of natural killer cells.
 A/Reference number: A47328; MUID:93133824; PMID:8421688
 A/Accession: B47328
 A>Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1507 <AND>
 A/Cross-references: GB:I04289; NID:9192866
 A/Note: authors translated the codon AGT for residue 972 as Arg
 R/Rintret, A.; Anderson, S.K. Mol. Immunol. 30, 1307-1313, 1993
 A>Title: IL-2 regulates the expression of the NK-TR gene via an alternate RNA splicing mechanism
 A/Reference number: I57820; MUID:94019422; PMID:8413330
 A/Accession: I77662
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 235-237/263-294 <RIN>
 A/Cross-references: GB:565998; NID:9425701; PIDN:AA028500.1; PID:9425702
 C/Genetics:
 A:Gene: NK-TR
 C/Superfamily: natural killer cell tumor-recognition protein; cyclophilin homology
 C/Keywords: alternative splicing; lymphocyte
 F/60-230/Domain: cyclophilin homology <CYP>

Query Match 48.9%; Score 43; DB 2; Length 1507;
 Best Local Similarity 81.8%; Pred. No. 64;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 RSRGSGRSGS 19
 DB 736 RSSGSGRSGS 746

Search completed: September 28, 2004, 06:15:02
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-8
Perfect score: 88
Sequence: 1 SHOE2XGRSGXSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	88.6	416	1	FILA_HUMAN
2	52	59.1	506	1	VE2_HPV47
3	46.5	52.8	135	1	RBE1_DROME
4	44	50.0	638	1	K220_HUMAN
5	43	48.9	694	1	FRQ2_DROME
6	43	48.9	1015	1	FRQ2_DROME
7	43	48.9	1453	1	NCROR_MOUSE
8	42	47.7	176	1	SSB_TREPA
9	42	47.7	772	1	MR11_CAEEL
10	41	46.6	825	1	SE5_FAT
11	41	46.6	843	1	DPO1_HPV47
12	41	46.6	983	1	Y144_HUMAN
13	40.5	46.0	1115	1	TBC2_CHLRE
14	40	45.5	150	1	RL19_PYRHO
15	40	45.5	151	1	RL19_PYRHO
16	40	45.5	330	1	DUSB_HUMAN
17	40	45.5	419	1	TCF7_MOUSE
18	40	45.5	526	1	FUS_HUMAN
19	40	45.5	660	1	DDXY_HUMAN
20	40	45.5	730	1	DPO1_HPV47
21	40	45.5	842	1	DPO1_HPV47
22	40	45.5	854	1	XRS2_YEAST
23	40	45.5	1232	1	YK05_CAEEL
24	39	44.3	526	1	YK05_CAEEL
25	39	44.3	617	1	DBP1_YEAST
26	39	44.3	786	1	DBP1_YEAST
27	39	44.3	713	1	DDX4_RAT
28	39	44.3	824	1	TGM1_RAT
29	39	44.3	834	1	YK22_YEAST
30	39	44.3	1337	1	DEXT_STROD
31	38.5	43.8	195	1	GRP3_DAUCA
32	38	43.2	124	1	SMO1_DROME
33	38	43.2	315	1	SOL2_YEAST

ALIGNMENTS

34	38	43.2	427	1	MYCN_SERCA	P26014 serinus can
35	38	43.2	502	1	VE2_HPV25	P36787 human papil
36	38	43.2	511	1	C7C4_ARATH	O64635 arabidopsis
37	38	43.2	648	1	FXR1_XENLA	P51115 xenopus lae
38	38	43.2	651	1	SEC9_YEAST	P40357 saccharomyc
39	38	43.2	1383	1	PRAX_RAT	O63425 rattus norv
40	38	43.2	3172	1	ERY1_SACER	O03133 saccharopol
41	38	43.2	3491	1	ERY1_SACER	O03131 saccharopol
42	37.5	42.6	857	1	V2A_CMYAT	O40977 cucumbei no
43	37.5	42.6	1054	1	IF2_STIAU	P55875 stigmateila
44	37	42.0	123	1	LSM4_CAEEL	O19952 caenorhabdi
45	37	42.0	168	1	DH1_MAIZE	P12950 zea mays (m

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
ID P26014: MYCN_SERCA
AC P26014: VE2_HPV25
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN FLC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69296901; PubMed=2740331;
RA McIntirey-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Ilesin S.R., Steinert P.M.,
RT "Characterization of a cDNA clone encoding human Filaggrin and
RT localization of the gene to chromosome region 1q21."
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RN [2]
RP CITRULLINATION.
RX MEDLINE=96374388; PubMed=8780679;
RA Sershan T., Kan S., Ogawa H., Marabe M., Asaga H.,
RT "Preferential delamination of keratin Ki and filaggrin during the
RT terminal differentiation of human epidermis."
RL Blocham. Biophys. Res. Commun. 225:712-719(1996).
CC -1- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -1- PFM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 34 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -1- PFM: Undergoes delamination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M24355; AAA52454.1; -
CC PIR; A32947; A32947.
CC GeneW: HGNC:3748; FLG.
CC MIW: 135940; -
CC GO; GO:0005482; C:intermediate filament; NAS.
CC GO; GO:0005188; F:structural molecule activity; NAS.
CC GO; GO:0007275; P:development; NAS.
CC InterPro; IPR003303; Filaggrin.

```

DR Pfam; PF03516; Filaggrin; 2.
DR PRINTS; PR00487; FILAGRIN.
KM Phosphorylation; Citrullination; Developmental protein.
FT NON TER
SO SEQUENCE 416 AA; 44105 MW; DEEA3218BA04332 CRC64;

Query Match 88.6%; Score 78; DB 1; Length 416;
Best Local Similarity 84.2%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRGRGSGRSGS 19
Db 7 SHOESTGRGRGSGRSGS 25

RESULT 2
ID V2 HPV47 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
"Genome organization and taxonomic position of human papillomavirus
RT type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE EAR RESPONSE ELEMENT (5'-ACCONNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2R'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -.
DR PIR; D35324; W2WL47.
DR HSSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C_1.
DR Pfam; PF00508; E2_N_1.
DR ProDom; PD000672; E2_C_1.
DR ProDom; PD000678; E2_N_1.
KM Early protein; Transcription regulation; Activator; DNA-binding;
KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SO SEQUENCE 506 AA; 57478 MW; 92C37F4BF8725065 CRC64;

Query Match 59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.27;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTGRGRGSGRSGS 19

```

```

Db 342 SRENTGRGRGSGRSGS 360

RESULT 3
ID RBPI DROME STANDARD; PRT; 135 AA.
AC Q02427; Q26271; Q9YGM8.
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein 1.
DE RBPI OR CG17136.
GN RBPI OR CG17136.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=94040720; PubMed=1340470;
RA Kim Y.-J., Zuo P., Manley J.L., Baker B.S.;
"The Drosophila RNA-binding protein RBPI is localized to
RT transcriptionally active sites of chromosomes and shows a functional
RT similarity to human splicing factor ASF/SF2."
RL Genes Dev. 6:2569-2579(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA Abiti J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley S.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintiner S.,
RA Borkova D., Botchan M.R., Bouck J., Brostein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Ciesla S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios R., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertire S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,
RA Palazolo M., Plittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Split B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yen R.-F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 14-53 FROM N.A.
RX MEDLINE=93109300; PubMed=8417324;
RA Kim Y.-J., Baker B.S.;
RT "Isolation of RRM-type RNA-binding protein genes and the analysis of

```

```

RT their relatedness by using a numerical approach."
RL Mol. Cell. Biol. 13:174-183(1993).
RN [4]
RP FUNCTION.
RX MEDLINE=95393975; PubMed=7664738;
RA Heinrichs V., Baker B.S.;
RT "The Drosophila SR protein RBPl contributes to the regulation of
RT doublesex alternative splicing by recognizing RBPl RNA target
RT sequences."
RL EMBL J. 14:3987-4000(1995).
CC -1- FUNCTION: Contributes to the activation of female-specific DSX
CC splicing in vivo by recognizing the RBPl target sequences within
CC the purine-rich polypyrimidine tract of the female-specific 3'
CC splice site.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=RBPl-A;
CC IsoId=Q02427-1; Sequence=Displayed;
CC Name=2; Synonyms=RBPl-B;
CC IsoId=Q02427-2; Sequence=VSP_005817;
CC -1- TISSUE SPECIFICITY: Ubiquitous
CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
CC -1- PTM: Extensively phosphorylated on serine residues in the RS
CC domain (Probable).
CC -1- SIMILARITY: Belongs to the splicing factor SR family.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04929; AAA28850.1;
DR EMBL: AE003688; AAF5455.1;
DR EMBL: S51691; AAB24622.1;
DR PIR: A46398; A46398.
DR PIR: A48110; A48110.
DR FlyBase: FBgn0010252; Rbpl.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PFO0076; rrm_1.
DR PROSITE: PS50102; RRM_1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KM Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
KM Phosphorylation.
FT DOMAIN 11 84 RNA-BINDING (RRM).
FT DOMAIN 82 135 ARG/SER-RICH (RS DOMAIN).
FT VARSPIC 107 135 Missing (in isoform 2).
FT CONFLICT 14 14 /FTID=VSP_005817.
FT CONFLICT 65 65 Y -> F (IN REF. 3).
FT CONFLICT 65 65 R -> A (IN REF. 1).
SQ SEQUENCE 135 AA; 15446 MW; 896DCE902518D991 CRC64;

Query Match 52.8%; Score 46.5; DB 1; Length 135;
Best Local Similarity 52.8%; Pred. No. 0.53;
Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 4 ESTXGRSR-----GXSGRSGS 19
DB 80 EMSGGRSDRRRRRGSGSSGSGS 102

RESULT 4
K220 HUMAN STANDARD; PRT; 638 AA.
AC 001546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2P) (K2P)

```

```

DE (CK 2P).
GN KRT2P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93114504; PubMed=1282112;
RA Collin C., Ouhayoun J.P., Grund C., Franke W.W.;
RT "Suprabasal marker proteins distinguishing keratinizing squamous
RT epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium
RT and epidermis are different."
RL Differentiation 51:137-148(1992).
CC -1- FUNCTION: Probably contributes to terminal cornification.
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC keratinocytes.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC (neutral to basic; 56-70 kDa) [K1 to K8].
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M99063; AAA35746.1;
DR GO: GO:0005198; C:intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
DR InterPro: IPR001664; IF.
DR InterPro: IPR002857; Keratin_I.
DR InterPro: IPR003054; Keratin_II.
DR Pfam: PFO0038; filament_1.
DR PRINTS: PRO1248; TYPE1KERATIN.
DR PRINTS: PRO1276; TYPE2KERATIN.
DR PROSITE: PS00226; IF_1.
KM Intermediate filament; Coiled coil; Keratin; Phosphorylation.
FT DOMAIN 1 182 HEAD.
FT DOMAIN 183 492 ROD.
FT DOMAIN 493 638 TAIL.
FT DOMAIN 183 218 COIL 1A.
FT DOMAIN 219 237 LINKER 1.
FT DOMAIN 238 329 COIL 1B.
FT DOMAIN 330 353 LINKER 12.
FT DOMAIN 354 492 COIL 2.
FT MOD RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 638 AA; 65871 MW; 9B743AB8B72076AF CRC64;

Query Match 50.0%; Score 44; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 7.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGSGS 17
DB 7 KXSFGSRGSGSGSGS 21

RESULT 5
FRZ2_DROME STANDARD; PRT; 694 AA.
AC 09VYX3; Q94916; Q9VYX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Frizzled protein 2 precursor (Frizzled-2) (dfzz).
DE FZ2 OR CG9739/CG14083.

```

CC Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhancu P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.;
 RT "A new member of the frizzled family from Drosophila functions as a
 RT WINGLESS receptor.";
 RL Nature 382:225-230 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731137;
 RA Adams M.D., Celtniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Baer A., Bakendale U., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintsov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Cantler A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Daveport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Nattai B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard D., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins. Inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeletons of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage

CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC -1- DOMAIN: Lys-Thr-X-X-X-Tyr motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
 CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL, U65589; AAC47273.1; -;
 CC DR EMBL, AE003518; AAF49185.2; -;
 CC DR F1R: 571786; S71786.
 CC DR FlyBase, FBgn0016797; fz2.
 CC DR GO: GO:0016021; C: integral to membrane, NAS.
 CC DR GO: GO:0017147; F: Wnt-protein binding, IDA.
 CC DR GO: GO:0006928; P: cell motility, IMP.
 CC DR GO: GO:0007163; P: establishment and/or maintenance of cell po. . .; ISS.
 CC DR GO: GO:0008585; P: female gonad development, IMP.
 CC DR GO: GO:0016055; P: Wnt receptor signaling pathway, IDA.
 CC DR InterPro: IPR000539; Frizzled.
 CC DR InterPro: IPR000024; Fz domain.
 CC DR InterPro: IPR000832; GPCR secretin.
 CC DR Pfam: PF01534; Frizzled_1.
 CC DR Pfam: PF01392; Fz_1.
 CC DR PRINTS: PR00489; FRIZZLED.
 CC DR SMART: SM00063; FRI: 1.
 CC DR PROSITE: PS50038; Fz: 1.
 CC DR PROSITE: PS50261; G_PROTEIN_RECEP_F2_4; 1.
 CC KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
 CC Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 CC FT SIGNAL 1 22
 CC FT CHAIN 23 694
 CC FT DOMAIN 23 315
 CC FT TRANSMEM 316 336
 CC FT DOMAIN 337 352
 CC FT TRANSMEM 353 373
 CC FT DOMAIN 374 397
 CC FT TRANSMEM 398 418
 CC FT DOMAIN 419 439
 CC FT TRANSMEM 440 460
 CC FT DOMAIN 461 482
 CC FT TRANSMEM 483 503
 CC FT DOMAIN 504 534
 CC FT TRANSMEM 535 555
 CC FT DOMAIN 556 584
 CC FT TRANSMEM 585 605
 CC FT DOMAIN 606 694
 CC FT TRANSMEM 187 225
 CC FT DOMAIN 59 180
 CC FT SITE 608 613
 CC FT SITE 692 694
 CC FT CAROXYD 78 78
 CC FT CAROXYD 288 288
 CC FT CONFLICT 55 55
 CC FT CONFLICT 417 417
 CC FT CONFLICT 417 417
 CC SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAB0956 CRC64;
 CC -----
 CC Query Match 48.9%; Score 43; DB 1; Length 694;
 CC Best Local Similarity 52.6%; Pred. No. 12;
 CC Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

QY      1 SHOESTXGRSGRSGSGS 19
DB      182 SYTEAGSGSGSGSGSGS 200

RESULT 6
FRQ_CRESP
ID      PRO_CRESP      STANDARD;      PRT; 1015 AA.
AC      000586;
DT      15-JUL-1999 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Frequency clock protein.
FRQ.
OS      Crepus spinulosus (Chromocrea spinulosa).
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX      NCBI_TaxId=110619;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FGSC 4896;
RX      MEDLINE=97188515; PubMed=90377100;
RA      Lewis M.T., Morgan L.W., Feldman J.F.;
RT      "Analysis of frequency (frq) clock gene homologs: evidence for a
RT      helix-turn-helix transcription factor.";
RL      Mol. Gen. Genet. 253:401-414(1997).
CC      -!- FUNCTION: Circadian clock component involved in the generation of
CC      biological rhythms, in particular in rhythm stability, period
CC      length, and temperature compensation. Behaves as a negative
CC      element in circadian transcriptional loop (by similarity).
CC      -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U25850; AAA68072.1; -
DR      PIR: T42013; T42013.
KW      Biological rhythms; Transcription regulation; Nuclear protein.
KW      DOMAIN
FT      DOMAIN 240 245
FT      DOMAIN 356 368
FT      DOMAIN 443 451
FT      DOMAIN 584 588
FT      DOMAIN 892 913
SQ      SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;

Query Match      48.9%; Score 43; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9313824; PubMed=8421688;
RA      Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
RA      Ortaido J.R.;
RT      "A cyclophilin-related protein involved in the function of natural
RT      killer cells.";
RL      Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RN      [2]
RP      REVISIONS TO C-TERMINUS.
RC      STRAIN=BALB/c; TISSUE=Blood;
RA      Anderson S.K.;
RT      Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: Component of a putative tumor-recognition complex.
CC      involved in the function of NK cells.
CC      -!- SIMILARITY: Contains 1 cyclophilin-like P1ase domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L04289; AAA37500.2; ALT_INIT.
DR      HSSP: Q27450; 1A33.
DR      MCD: MGI:97346; Nktr.
DR      InterPro: IPR002130; CSA_P1ase.
DR      Pfam: PF00160; PRO_Isomerase; 1.
DR      PRINTS: PRO0153; CSA_P1ase.
DR      PROSITE: PS00170; CSA_P1ase_1; 1.
DR      PROSITE: PS50072; CSA_P1ase_2; 1.
KW      Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
KW      DOMAIN
FT      DOMAIN 1 176
FT      DOMAIN 222 241
FT      DOMAIN 422 459
FT      DOMAIN 964 1003
FT      DOMAIN 198 273
FT      DOMAIN 468 565
FT      DOMAIN 658 812
FT      DOMAIN 1303 1453
SQ      SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;

Query Match      48.9%; Score 43; DB 1; Length 1453;
Best Local Similarity 81.8%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 SHOESTXGRSGRSGSGS 19
DB      583 NHRKXGTGSHSTGDSGSGSN 601

RESULT 7
NKR_MOUSE
ID      NKR_MOUSE      STANDARD;      PRT; 1453 AA.
AC      P30415;
DT      01-APR-1993 (Rel. 25, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      NK-tumor recognition protein (Natural-killer cells cyclophilin-
DE      related protein) (NKR-TR protein).
GN      NKRTR.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;

```

```

QY      9 RSRGXSGRSGS 19
DB      682 RSSGSGRSGS 692

RESULT 8
SSB_TREPA
ID      SSB_TREPA      STANDARD;      PRT; 176 AA.
AC      O83101;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Single-strand binding protein (SSB) (Helix-destabilizing protein).
DE      SSB OR TP0062.
GN      Treponema pallidum.
OS      Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OC      Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX      NCBI_TaxId=160;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Nichols;
RX      Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RX      Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA      Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA      Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,

```

```

RA McDonald L., Arriach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -1- FUNCTION: This protein is essential for replication of the
CC chromosome. It is also involved in DNA recombination and repair
CC (by similarity).
CC -1- SIMILARITY: Contains 1 SSB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001191; AAC65057.1; -.
DR PIR; F71370; F71370.
DR HSSP; P02339; IEYG.
DR TIGR; TP0062; -.
DR InterPro; IPR008994; Nucleic_acid_DB.
DR InterPro; IPR000424; SSB_protein.
DR Pfam; PF00436; SSB; 1.
DR TIGRFAMs; TIGR00621; ssb; 1.
DR PROSITE; PS50935; SSB; 1.
DR DNA-binding; DNA repair; DNA replication; Complete proteome.
KW DOMAIN
FT 4 106 SSB.
SQ SEQUENCE 176 AA; 19043 MW; 47C49D2F79927D2C CRC64;

Query Match 47.7%; Score 42; DB 1; Length 176;
Best Local Similarity 62.5%; Pred. No. 4;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ESTXGRGRGXSGRSGS 19
Db 128 ESTGRVGTSTSDERS 143

RESULT 9
MR11_CAEEL STANDARD; PRT; 772 AA.
ID MR11_CAEEL
AC Q23255;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Double-strand break repair protein mre-11.
GN MRE-11 OR ZC302.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP G10-139.
RX MEDLINE=21135651; PubMed=11238374;
RA Chin G.M., Villeneuve A.M.;
RT "C. elegans mre-11 is required for meiotic recombination and DNA
RT repair but is dispensable for the meiotic G(2) DNA damage
RT checkpoint."
RL Genes Dev. 15:522-534(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kelly P.F.;
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
CC Possesses single-strand endonuclease activity and double-strand-
CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
CC processing.
CC -1- FUNCTION: Required for meiotic crossing over and chiasma

```

```

CC formation. Pachytene morphology and homolog pairing are normal.
CC Vital in long term for maintenance of reproductive capacity of
CC subsequent generations.
CC -1- COPACITOR: Manganese (By similarity).
CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the MRE11/RAD32 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73978; CAA8292.1; -.
DR PIR; T27512; T27512.
DR GenBank; ZC302.1; C06573.
DR WormPep; ZC302.1; C06573.
DR InterPro; IPR003701; DNA_repair.
DR InterPro; IPR004643; M-peptidase.
DR InterPro; IPR007281; Mre11_DNA_bind.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF04152; Mre11_DNA_bind; 1.
DR TIGRFAMs; TIGR00583; mre11; 1.
DR DNA_repair; Hydrolase; Nuclease; Endonuclease; Exonuclease;
KW Nuclear protein; Manganese; Meiosis.
FT 139 139 E-K; IN MRE-11-ME41; DEFECTIVE IN
FT MUTAGEN
SQ SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;

Query Match 47.7%; Score 42; DB 1; Length 772;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGXSGRSGS 19
Db 697 SKQPTGRGRGRGARGAGAS 715

RESULT 10
SES_RAT STANDARD; PRT; 825 AA.
ID SES_RAT
AC Q63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SES antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Miscart; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein SES in the nervous system."
RL J. Biochem. 118:122-128(1995).
RN [2]
RP FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: D37934; BAA07153.1; -
 DR PIR: JG4163; JG4163.
 KM DNA-binding; Nuclear protein; Antigen.
 SQ SEQUENCE 825 AA; 86831 MW; AF667FE2PD55BDF CRC64;
 Query Match 46.6%; Score 41; DB 1; Length 825;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 8 GRSRGSGRGS 19
 DB 589 GRGRGGRGGRGS 600
 RESULT 11
 ID DPOL_HPBVR STANDARD; PRT; 843 AA.
 AC P03157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN Hepatitis B virus (subtype adr).
 OS Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_Taxid=106820;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=83168919; PubMed=6300776;
 RX Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw.";
 RL Nucleic Acids Res. 11:1747-1757(1983).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA)(N).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-phosphomonoester.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: V00867; -1 NOT_ANNOTATED_CDS.
 DR PIR: A00704; JDIYVR.
 DR INTERPRO: IPR001462; DNAPOL_viral_C.
 DR INTERPRO: IPR000201; DNAPOL_viral_N.
 DR INTERPRO: IPR000477; RVase.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; rvc2.
 DR Pfam: PF000814; DNAPOL_viral_C; 1.
 DR Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase; Hydrolase; Nuclease; Endonuclease; DNA replication; DNA-binding.
 KM SQUENCE 843 AA; 94400 MW; A6B2D49083C458B CRC64;
 SQ
 Query Match 46.6%; Score 41; DB 1; Length 843;
 Best Local Similarity 52.8%; Pred. No. 31;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 3 QESTXGRSGRGS 19
 DB 226 QPQGGSWARGSGRGS 242
 RESULT 12
 ID Y144_HUMAN STANDARD; PRT; 993 AA.
 AC Q14157;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein KIAA0144.
 GN KIAA0144.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:167-174(1995).
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: D63478; BAA09765.1; -
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam: PF00627; UBA; 1.
 DR SMART: SM00165; UBA; 1.
 DR PROSITE: PS00030; UBA; 1.
 KM Hypothetical protein.
 FT DOMAIN 49 89 UBA.
 SQ SEQUENCE 983 AA; 103930 MW; 50579CFA2931EBA68 CRC64;
 QY 8 GRSRGSGRGS 18
 DB 178 GRGRGSGRGRG 188
 RESULT 13
 ID TBCC2_CHLRE STANDARD; PRT; 1115 AA.
 AC Q8VXF3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE TBCC2 translation factor, chloroplast precursor.
 GN TBCC2.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonadaceae;
 OX NCBI_Taxid=3055;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=137C / CC-125;
 RX MEDLINE=22053252; PubMed=12045185;
 RA Auchincloss A.H., Zerges W., Perron K., Girard-Bascou J., Rochaix J.-D.;
 RT "Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbc mRNA in Chlamydomonas reinhardtii.";
 RT RT required for translation of the chloroplast psbc mRNA in Chlamydomonas reinhardtii.";
 RL Cell Biol. 157:953-962(2002).
 CC -1- FUNCTION: Required for expression of the chloroplast encoded psbc mRNA, most likely for translation initiation. Interacts with the 5'UTR of psbc.
 CC -1- SUBUNIT: Part of a 400 kDa complex which is not stably associated


```

CC      with RNA.
CC      -1- SIMILARITY: IN THE CENTRAL SECTION. TO CR1 OF ZEA MAYS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ27966; CAP20887.1; -.
KW      Chloroplast; Transalt peptide; Repeat.
FT      TRANSLAT 1
FT      CHAIN 1
FT      DOMAIN 483 1115 CHLOROPLAST (POTENTIAL).
FT      REPEAT 483 1068 TBC2 TRANSLATION FACTOR.
FT      REPEAT 607 521 9 X 38 AA APPROXIMATE REPEATS.
FT      REPEAT 685 723 1.
FT      REPEAT 724 763 2.
FT      REPEAT 764 803 3.
FT      REPEAT 804 842 4.
FT      REPEAT 843 880 5.
FT      REPEAT 990 1029 6.
FT      REPEAT 1030 1068 7.
FT      DOMAIN 187 210 SER-RICH.
FT      DOMAIN 253 256 POLY-ASP.
FT      DOMAIN 271 326 POLY-RICH.
FT      DOMAIN 370 377 POLY-GLN.
FT      DOMAIN 883 1115 ALA-RICH.
SQ      SEQUENCE 1115 AA; 114823 MW; 2239799B91C5D8F7 CRC64;

Query Match 46.0%; Score 40.5; DB 1; Length 1115;
Best Local Similarity 55.6%; Pred. No. 51;
Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY      5 STXGRSG---XSGRSGS 19
DB      170 STSGRAGMGSGGRNGS 187

RESULT 14
RL19 PYRHO STANDARD; PRT; 150 AA.
AC      OS9437;
DT      15-DEC-1998 (Rel. 37, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DE      50S ribosomal protein L19E.
GN      RPL19E OR PH1759.
OS      Pyrococcus horikoshii.
OC      Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC      Pyrococcus.
CC      NCBI_TaxID=53953;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=OT3;
RX      MEDLINE=98344137; Pubmed=9679194;
RA      Kawarayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA      Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA      Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtoku Y.,
RA      Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kushiida N., Oguchi A.,
RA      Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA      Masuchi Y., Shizuya H., Kikuchi H.;
RT      "Complete sequence and gene organization of the genome of a hyper-
RT      thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL      DNA Res. 5:55-76(1998).
CC      -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AP000007; BAA30873.1; -.
DR      PIR; B71185; B71185.
DR      InterPro; IPR000196; Ribosomal_L19e.
DR      Pfam; PF01280; Ribosomal_L19e; 1.
DR      ProDom; PD004823; Ribosomal_L19e; 1.
DR      PROSITE; PS00526; RIBOSOMAL_L19E; 1.
KW      Ribosomal protein; Complete proteome.
SQ      SEQUENCE 150 AA; 1797 MW; 41223A854825B69 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 150;
Best Local Similarity 41.2%; Pred. No. 7.2;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      2 HOESTXGRSGXSGRSG 18
DB      67 HEQKKGRHGRGSGRKG 83

RESULT 15
RL19 PYRAB STANDARD; PRT; 151 AA.
AC      Q9V1V3;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      50S ribosomal protein L19E.
GN      RPL19B OR PYRAB01230 OR PAB2134.
OS      Pyrococcus abyssi.
OC      Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC      Pyrococcus.
CC      NCBI_TaxID=29292;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=GS / Orsay;
RX      MEDLINE=22511545; Pubmed=12622808;
RA      Cohen G.N., Barde V., Flament D., Galperin M., Helling R., Lecompte O.,
RA      Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA      Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT      "An integrated analysis of the genome of the hyperthermophilic
RT      archaeon Pyrococcus abyssi."
RL      Mol. Microbiol. 47:1495-1512(2003).
CC      -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ248284; CAB49245.1; -.
DR      PIR; F75145; F75145.
DR      InterPro; IPR000196; Ribosomal_L19e.
DR      Pfam; PF01280; Ribosomal_L19e; 1.
DR      ProDom; PD004823; Ribosomal_L19e; 1.
DR      PROSITE; PS00526; RIBOSOMAL_L19E; 1.
KW      Ribosomal protein; Complete proteome.
SQ      SEQUENCE 151 AA; 18187 MW; 85FC805644B437A0 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 151;
Best Local Similarity 41.2%; Pred. No. 7.2;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      2 HOESTXGRSGXSGRSG 18
DB      67 HEQKKGRHGRGSGRKG 83

```


Tue Sep 28 07:11:39 2004

us-09-308-150-8.open.rsp

Page 9

Search completed: September 28, 2004, 06:05:51
Job time : 7.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	84	95.5	797	4	Q16824 homo sapien
2	84	95.5	990	4	Q15206 homo sapien
3	84	95.5	1218	4	Q05331 homo sapien
4	78	86.6	798	4	Q9H4U3 homo sapien
5	78	88.6	1084	4	Q01212 homo sapien
6	69	78.4	591	4	Q01720 homo sapien
7	69	78.4	687	4	Q9H4U2 homo sapien
8	64	72.7	465	4	Q03838 homo sapien
9	63	71.6	322	4	Q075370 homo sapien
10	49	55.7	822	3	Q9P312 neurospora
11	46.5	52.8	144	5	Q8T9K6 arabidopsis
12	46	52.3	1142	13	Q7S2V2 xenopus lae
13	45	51.1	2464	10	Q9SL90 arabidopsis
14	44	50.0	285	5	Q9M3V3 drosophila
15	44	50.0	638	4	Q7Z795 homo sapien
16	44	50.0	652	5	Q77168 apis mellif

17	44	50.0	800	5	Q15845 leishmania
18	44	50.0	936	16	Q82H71 streptomyce
19	44	50.0	1284	5	Q9V9Y3 drosophila
20	44	50.0	1829	16	Q96808 streptomyce
21	43	48.9	185	10	Q8LNN6 oryza sativ
22	43	48.9	306	5	Q45386 caenorhabd
23	43	48.9	448	4	Q8NET3 gnet3 homo sapien
24	43	48.9	452	4	Q96PY4 gnet3 homo sapien
25	43	48.9	820	12	Q69375 mouse cytom
26	43	48.9	849	10	Q9C7M2 arabidopsis
27	43	48.9	849	10	Q949M6 arabidopsis
28	43	48.9	966	11	Q8CBP6 mus musculu
29	42.5	48.3	240	16	Q87YL4 pseudomonas
30	42	47.7	259	4	Q86YA2 homo sapien
31	42	47.7	322	16	Q8G4B5 bifidobacte
32	42	47.7	481	16	Q7UJ99 rhodospirill
33	42	47.7	709	17	Q8T9F1 methanosarc
34	42	47.7	983	11	Q8B0S3 mus musculu
35	42	47.7	983	11	Q812D4 mus musculu
36	42	47.7	1014	11	Q8BIT6 mus musculu
37	42	47.7	1015	11	Q8BIW4 mus musculu
38	42	47.7	1035	11	Q812D5 mus musculu
39	42	47.7	1067	11	Q8CIG7 mus musculu
40	42	47.7	1084	13	Q7S1B7 brachydanio
41	42	47.7	1105	11	Q8K102 mus musculu
42	42	47.7	1107	11	Q80X50 mus musculu
43	42	47.7	1112	11	Q8B701 mus musculu
44	42	47.7	1229	3	Q8B2H1 mus musculu
45	41.5	47.2	958	5	Q9VC54 drosophila

ALIGNMENTS

RESULT 1
Q16824 PRELIMINARY; PRT; 797 AA.

AC Q16824; 01-NOV-1996 (TRENBLER. 01, Created)
DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLER. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:91064347; PubMed:2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinhart P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin gene [published erratum appears in Biochemistry 1991 Jun 11;30(23):5814].";
RT 11,30(23):5814.
RL Biochemistry 29:9432-9440(1990).
DR EMBL: M60502; AAA63248.1; -.
DR GO: GO:0005198; Functional molecule activity; IEA.
DR InterPro: IPR003303; Profilaggrin.
DR Pfam: PF03516; Filaggrin; 4.
DR PRINTS: PR00487; FILAGGRIN.
FT NON TER 1
SQ SEQUENCE 797 AA; 85176 MW; 60B6184763BDA86B CRC64;

Query Match 95.5%; Score 84; DB 4; Length 797;
Best Local Similarity 89.5%; Pred. No. 8e-06; 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQESTXGRSGRSGSGS 19
Db 427 SHQESTRGRSGRSGSGS 445

RESULT 2

Q15206 PRELIMINARY; PRT; 990 AA.
 ID Q15206
 AC Q15206;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 GN GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinhert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]";
 RL Biochemistry 29:9432-9440(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinhert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene";
 RL Biochemistry 30:5814-5814(1991).
 DR EMBL; M60494; AAA63244.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Flaggrin.
 DR Pfam; PF03516; Flaggrin; 6.
 DR PRINTS; PR00487; FLAGGRIN.
 DR NON TER 990
 FT SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
 SQ
 Query Match 95.5%; Score 84; DB 4; Length 990;
 Best Local Similarity 89.5%; Pred. No. 1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 Db 227 SHQSTXGRSGRSGSGS 245
 RESULT 3
 ID Q05331 PRELIMINARY; PRT; 1218 AA.
 AC Q05331;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE FLAGGRIN (PROFILAGGRIN) (Fragment).
 GN GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FORESKIN;
 RX MEDLINE=93109348; PubMed=8417356;
 RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
 RA Steinhert P.M.;
 RT "Profilaggrin is a major epidermal calcium-binding protein.";
 RL Mol. Cell. Biol. 13:613-625(1993).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS
 CC -1- PFM: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
 CC REPEATS.
 DR EMBL; M69943; AAA64487.1; -;
 DR PIR; A48118; A48118.
 DR HSSP; P02593; 1CDM.
 DR GO; GO:0005856; C:cycloskeleton; NAS.
 DR GO; GO:0005509; P:calcium ion binding; TAS.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
 DR InterPro; IPR001751; Cabp_S100.
 DR InterPro; IPR002048; Bf-hand.
 DR InterPro; IPR003303; Flaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Flaggrin; 6.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FLAGGRIN.
 DR PROSITE; PS00018; Bf_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KW Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
 KW Polymorphism.
 FT CA_BIND 19 32 SITE I (BY SIMILARITY).
 FT CA_BIND 62 73 SITE II (BY SIMILARITY).
 FT NON TER 1218 1218
 SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5265B19C2 CRC64;
 Query Match 95.5%; Score 84; DB 4; Length 1218;
 Best Local Similarity 89.5%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 Db 449 SHQSTXGRSGRSGSGS 467
 RESULT 4
 ID Q9H4U3 PRELIMINARY; PRT; 798 AA.
 AC Q9H4U3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE DJL1N1.1.2 (Profilaggrin 3' end) (Fragment).
 GN GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL356504; CAC3171.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Flaggrin.
 DR Pfam; PF03516; Flaggrin; 4.
 DR PRINTS; PR00487; FLAGGRIN.
 DR NON TER 1
 FT SEQUENCE 798 AA; 84773 MW; P923DDAD8D1290805 CRC64;
 SQ
 Query Match 88.6%; Score 78; DB 4; Length 798;
 Best Local Similarity 84.2%; Pred. No. 8.6e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 Db 428 SHQSTXGRSGRSGSGS 446
 RESULT 5
 ID Q01212 PRELIMINARY; PRT; 1084 AA.
 CC Q01212

```

AC 001212; 003840;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; JOINED.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; P:development; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin.
DR PRINTS; PRO0487; FILAGGRIN.
FT NON TER
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 88.6%; Score 78; DB 4; Length 1084;
Best Local Similarity 84.2%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 64 SHOESTXGRSGRSGSGS 82

RESULT 6
Q01720 PRELIMINARY; PRT; 591 AA.
AC 001720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Ntunusukit W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus.";
RL J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 34 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PRO0487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293 POTENTIAL.
FT CHAIN 294 467 FILAGGRIN.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 591 FILAGGRIN.
FT CA_BIND 19 32 SITE 1 (BY SIMILARITY).
FT CA_BIND 62 73 SITE 11 (BY SIMILARITY).
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.0022;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 449 SHOESTXGRSGRSGSGS 467

RESULT 7
Q094U2 PRELIMINARY; PRT; 687 AA.
AC 094U2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Laid G.;
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PRO0487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.0025;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 SHOESTXGRSGRSGSGS 19
DB 449 SHOESTRGRSGRSGSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003838;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE FILAGGRIN (PROFILAGGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

PC TISSUE=PLACENTA;

RX MEDLINE=91064347; PubMed=2248957;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 29:9432-9440(1990).

RN [2]

RP REVISIONS.

RX MEDLINE=91255199; PubMed=2043621;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISMS: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.

CC -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

DR EMBL; M60499; AAA63246.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 3.

DR PRINTS; PR00487; FILAGGRIN.

FT NON_TER 1

FT NON_TER 465

FT SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

Query Match 72.7%; Score 64; DB 4; Length 465;
Best Local Similarity 73.7%; Pred. No. 0.012;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 227 SHOESARGRSGRSGSGS 245

RESULT 9

ID 075370 PRELIMINARY; PRT; 322 AA.

AC 075370;

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Epidermal filaggrin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;

RA Giral-Neuhäuser E., Durieux J.J., Arnaud M., Dalbon P., Sebba M., Vincent C., Simon M., Sennu T., Masson-Bessière C., Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated anti-filaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deamination of arginine residues.";

RL J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC23559.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.

DR PRINTS; PR00487; FILAGGRIN.

FT NON_TER 1

FT NON_TER 322

FT SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match 71.6%; Score 63; DB 4; Length 322;
Best Local Similarity 72.2%; Pred. No. 0.012;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 18
DB 305 SHOESTRGRSGRSGSGS 322

RESULT 10

ID 099312 PRELIMINARY; PRT; 822 AA.

AC 099312;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Related to nucleolar phosphoprotein.

GN B12P1.10.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCBI_TaxID=5141;

RN [1]

RP SEQUENCE FROM N.A.

RA Schulte U., Aign V., Hehseisel U., Brandt P., Fartmann B., Holland R., Nyakatura G., Nemes H.W., Mannhaupt G.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA German Neurospora genome project;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL390091; CAB98213.1; -

DR PIR; T51049; T51049.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR000504; RNA_rec_mol.

DR InterPro; IPR005120; Smg-4_UFP3.

DR Pfam; PF00076; rrm; 1.

DR Pfam; PF03467; Smg4_UFP3; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS0102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

FT SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 55.7%; Score 49; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 8.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTXGRSGRSGSGS 19
DB 414 RESAGSTRGRSGRSGT 430

RESULT 11

08T9K6 PRELIMINARY; PRT; 144 AA.
 ID 08T9K6
 AC 08T9K6
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE RM02602D (CG17136-PD).
 GN Bp1 OR CG17136.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 (1)
 RA SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guanin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 (2)
 RA SEQUENCE FROM N.A.
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabo R.D.,
 RA Abil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Balaban D.,
 RA Bailes R.M., Bassu P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Eranenilata C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.S., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai Z.,
 RA Lasok P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen T.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski R., Smith T.,
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan G., Zhao Q., Zhao Q., Zhu L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of Drosophila melanogaster.",
 RA Science 287:2185-2195(2000).
 (3)
 RA SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Barzon J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Carlson J.W., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Matvei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (4)
 RA SEQUENCE FROM N.A.
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminler J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bertan B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,
 RA Seakle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RA "Annotation of Drosophila melanogaster genome.",
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (5)
 RA SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (6)
 RA SEQUENCE FROM N.A.
 RA Flybase;
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069252; AL139397.1; -;
 DR EMBL; AB003688; AAN13487.1; -;
 DR Flybase; PB0010252; Bp1.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR SEQUENCE 144 AA; 16013 MW; AB5BD782CED5A46 CRC64;
 SQ
 Query Match 52.8%; Score 46.5; DB 5; Length 144;
 Best Local Similarity 52.8%; Pred. No. 3.4;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
 4 ESTXGRSR-----GXSGRSGS 19
 Db 80 EMSGRSRDRRRGSGSGSGS 102
 RESULT 12
 Q7SZV2 PRELIMINARY; PRT; 1142 AA.
 ID Q7SZV2
 AC Q7SZV2
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Period 1.
 GN PER1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxId=8335;
 (1)
 RA SEQUENCE FROM N.A.
 RA Stancace C.M., Sutton B., Taylor M., Green C.B.,
 RA "Xenopus laevis period 1.",
 RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY340103; AAQ18143.1; -;
 DR SEQUENCE 1142 AA; 1580CB698D9E0B6 CRC64;
 SQ
 Query Match 52.3%; Score 46; DB 13; Length 1142;
 Best Local Similarity 58.8%; Pred. No. 38;
 Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 1 SHQESTXGRSRXSGRS 17
 Db 805 SKGSRSGRRRSGSGS 821

RESULT 13

Q9SL90 PRELIMINARY; PRT; 2464 AA.

ID Q9SL90

AC Q9SL90

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Hypothetical protein.

GN A1825730.

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

[1]

RA SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M., RA Shen M., Roming C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RM [2]

RA SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RM [3]

RA SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006053; AAD3133.2; -.

DR PIR; A84652; A84652.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR005797; Cytb_b6_N.

DR InterPro; IPR008941; TPR-like.

DR PROSITE; PS00192; CYTOCHROME_B_HEME_1.

DR Hypothetical protein.

SW SEQUENCE 2464 AA; 274770 MW; 63E0942007645BSP CRC64;

QY 5 STXGRSGXSGRSGS 19

Db 1353 STGRSGXGCGGGSNS 1367

Query Match

Best Local Similarity 51.1%; Score 45; DB 10; Length 2464;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RESULT 14

Q9W3V3 PRELIMINARY; PRT; 285 AA.

ID Q9W3V3

AC Q9W3V3

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE CCL4435 protein.

GN CCL4435.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RA SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D., RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X., RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D., RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., RA Ballew R.M., Basu A., Baxendale U., Bayraktoglu L., Beasley E.M., RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S., RA Borova D., Botchan M.R., Bouck U., Brokstein P., Brotlier P., RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., RA Glodek A., Gong F., Gorelli J.H., Gu Z., Guan P., Harris M., RA Harris N.L., Harvey D., Heiman J.U., Hernandez J.R., Houck J., RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C., RA Ualali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D., RA Mervlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacled J.M., RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., RA Switzkas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL EMBL; AE003439; AAF46210.1; -.

DR FLYbase; FBgn0029911; CG14435.

SW SEQUENCE 285 AA; 29822 MW; 95BDE3388D4B78D3 CRC64;

QY 2 HOESTXGRSGXSGRSG 18

Db 38 HOHSNPGQNRGNNGGCGG 54

Query Match

Best Local Similarity 50.0%; Score 44; DB 5; Length 285;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

RESULT 15

Q7Z795 PRELIMINARY; PRT; 638 AA.

ID Q7Z795

AC Q7Z795

DT 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Keratin 2p.

GN KRT2B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RA SEQUENCE FROM N.A.

RA Rogers W.A., Langbein L., Smith F.J.D., McLean W., Winter H., RA Beckmann I., Praetzel S., Schweizer J.,

RT "Characterisation of new members of the human Type II keratin gene family."

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ564103; CAD91891.1; -.

SW SEQUENCE 638 AA; 65840 MW; 6A312B8260BE06A9 CRC64;

QY 638 AA; 65840 MW; 6A312B8260BE06A9 CRC64;

Query Match

Best Local Similarity 50.0%; Score 44; DB 4; Length 638;

Tue Sep 28 07:11:40 2004

us-09-308-150-8.open.rsppt

Page 7

Best Local Similarity	60.0%	Pred. No. 45;			
Matches	9; Conservative	3; Mismatches	3; Indels	0; Gaps	0;

```
QY      3 QESTXGRSRGXSGRS 17
          ::|||:|||||
Db      7 KKSFGSRSGGFSGRS 21
```

```
Search completed: September 28, 2004, 06:12:47
Job time : 41.1771 secs
```

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-8

Sequence: 1 SHQESTXGRSGXGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: A_Geneseq_299and4:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
9: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	2	AAW61515
2	84	95.5	19	2	AAW61508
3	84	95.5	19	2	AAW61512
4	84	95.5	19	2	AAW61517
5	84	95.5	19	2	AAW61505
6	84	95.5	19	2	AAW61514
7	84	95.5	19	2	AAW61516
8	84	95.5	19	2	AAW61506
9	78	88.6	19	2	AAW61507
10	78	88.6	19	2	AAW61509
11	78	88.6	19	2	AAW61511
12	78	88.6	19	2	AAW61513
13	78	88.6	19	2	AAW61510
14	75	85.2	22	4	AAW61520
15	75	85.2	22	4	AAW61525
16	60	68.2	330	2	AAW22956
17	60	68.2	330	2	AAW22957
18	60	68.2	330	2	AAW22954
19	60	68.2	330	2	AAW22955
20	55	63.6	441	4	ABG19110
21	51	56.8	477	6	ABO07142
22	46.5	52.8	135	4	ABW6181
23	45	51.1	1711	4	AAW79819
24	45	51.1	1951	4	AAW78835
25	45	51.1	1951	4	AAW78835

26	44	50.0	285	4	ABB69359	Abb69359	Drosophila
27	44	50.0	1199	4	ABB58274	Abb58274	Drosophila
28	43	48.9	73	6	AAU54919	Aau54919	Propionib
29	43	48.9	73	6	AAU54919	Aau54919	Propionib
30	43	48.9	452	6	AAO16179	Aao16179	Human pro
31	43	48.9	569	4	AAW79339	Aaw79339	Human pro
32	43	48.9	569	4	AAW79340	Aaw79340	Human pro
33	43	48.9	694	2	AAW31267	Aaw31267	Drosophila
34	43	48.9	694	2	ABW71797	Abw71797	Drosophila
35	43	48.9	797	5	ABP73371	Abp73371	Candida a
36	43	48.9	982	2	AAW13320	Aaw13320	Murine Na
37	43	48.9	982	6	ABU05134	Abu05134	Human exp
38	43	48.9	1482	7	ADE58671	Ade58671	Rat Prote
39	43	48.9	1482	7	ADE58675	Ade58675	Rat Prote
40	42	47.7	176	6	ABU48456	Abu48456	Protein e
41	42	47.7	772	4	AAW64573	Aaw64573	Human Mre
42	41.5	47.2	957	4	ABB62605	Abb62605	Drosophila
43	41.5	47.2	1938	6	ABP76679	Abp76679	Streptomy
44	41	46.6	77	5	ABP06528	Abp06528	Human ORF
45	41	46.6	184	3	AAW57041	Aaw57041	Human pro

ALIGNMENTS

RESULT 1
AAW61515
ID AAW61515 standard; peptide: 19 AA.
AC AAW61515;
XX
XX
DT 26-OCT-1998 (first entry)
XX
XX
DE Peptide cfa, based on cDNA of a profilaggrin repeat.
XX
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX
PN WO9822503-A2.
XX
XX
PD 28-MAY-1998.
XX
XX
PF 14-NOV-1997; 97WO-NL000624.
XX
XX
PR 15-NOV-1996; 96NL-01004539.
XX
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
PI Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA,
XX
XX
DR WPI, 1998-398613/34.
XX
XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX
PS Disclosure; Page 6; 19PP; English.
XX
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.5e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

RESULT 2
 ID AAW61508 standard; peptide; 19 AA.

XX AAW61508;

XX 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX Solid phase synthesis; peptide amide; polyclonal antibody;

XX Monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX Location/Qualifiers

XX /note= "Citruilline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

AAW61512
 ID AAW61512 standard; peptide; 19 AA.

XX AAW61512;

XX 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX Solid phase synthesis; peptide amide; polyclonal antibody;

XX Monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX Location/Qualifiers

XX /note= "Citruilline"

XX Modified-site

XX /note= "Citruilline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

XX 26-OCT-1998 (first entry)

XX Peptide cfc0, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX Solid phase synthesis; peptide amide; polyclonal antibody;

XX Monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX Location/Qualifiers

XX /note= "Citruilline"

XX Modified-site

XX /note= "Citruilline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

RESULT 3

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

RESULT 4

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

AAW61517
 ID AAW61517 standard; peptide; 19 AA.
 XX AAW61517;
 XX 26-OCT-1998 (first entry)
 XX Peptide cfc0, based on cDNA of a profilaggrin repeat.

KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW Solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 OS Synthetic.
 OS Homo sapiens.
 XX MO9822503-A2.
 PN 28-MAY-1998.
 PD 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR Peptide derived from an antigen recognised by autoantibodies - is
 XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 SQ
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 8e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGXSGRSGS 19
 DB 1 SHQESTXGRSRGXSGRSGS 19
 RESULT 5
 AAW61505
 ID AAW61505 standard; peptide; 19 AA.
 XX
 AC AAW61505;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
 DS
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 7 /note= "Citruiline"
 FT
 XX
 PN MO9822503-A2.
 XX 28-MAY-1998.
 PD

PF 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR Peptide derived from an antigen recognised by autoantibodies - is
 XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 SQ
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGXSGRSGS 19
 DB 1 SHQESTXGRSRGXSGRSGS 19
 RESULT 6
 AAW61514
 ID AAW61514 standard; peptide; 19 AA.
 XX
 AC AAW61514;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cf, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX MO9822503-A2.
 PN 28-MAY-1998.
 PD 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR Peptide derived from an antigen recognised by autoantibodies - is
 XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from patients with rheumatoid arthritis (RA). This peptide is reactive with a RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which produced them as peptide amides. These sequences may be used in the detection of RA autoimmune antibodies, in the diagnosis of RA, as well as for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQSTXGRSRGSGRSGS 19
1 SHQSTXGRSRGSGRSGS 19

RESULT 7
AAW61516
ID AAW61516 standard; peptide; 19 AA.

XX AAW61516;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cFE, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Scheilekens GA, Raats JMF, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from patients with rheumatoid arthritis (RA). This peptide is reactive with a RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which produced them as peptide amides. These sequences may be used in the detection of RA autoimmune antibodies, in the diagnosis of RA, as well as for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQSTXGRSRGSGRSGS 19
1 SHQSTXGRSRGSGRSGS 19

RESULT 8
ABW97605
ID ABW97605 standard; protein; 1467 AA.

XX ABB97605;

XX 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 873.

XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX antiinfectivity; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.

XX Homo sapiens.

XX W0200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Weinman T, Drmanac RT;

XX WPI; 2002-292408/33.

XX N-PSDB; ABN32791.

XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 1467 AA;

Query Match 95.5%; Score 84; DB 5; Length 1467;
Best Local Similarity 89.5%; Pred. No. 5.6e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQSTXGRSRGSGRSGS 19
449 SHQSTXGRSRGSGRSGS 467

RESULT 9
AAW61506
ID AAW61506 standard; peptide; 19 AA.

XX AAW61506;

DT 26-OCT-1998 (first entry)
 XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSRGSGSGS 19
 DB 1 SHOESTYGRSRGSGSGS 19
 RESULT 10
 AAM61507
 ID AAM61507 standard; peptide; 19 AA.
 XX
 AC AAM61507;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX
 KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 11 /note= "Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSRGSGSGS 19
 DB 1 SHOESTYGRSRGSGSGS 19
 RESULT 11
 AAM61509
 ID AAM61509 standard; peptide; 19 AA.
 XX
 AC AAM61509;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 16 /note= "Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 7.6e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTYGRSRGSGSGS 19
 DB 1 SHQESTYGRSRGSGSGS 19
 RESULT 12
 AAW61511
 ID AAW61511 standard; peptide; 19 AA.
 XX
 AC AAW61511;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DS Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 7
 FT Modified-site /note= "Citruilline"
 FT Modified-site 11
 FT /note= "Citruilline"
 XX
 W09822503-A2.
 XX
 PN 28-MAY-1998.
 XX
 PD 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTYGRSRGSGSGS 19
 DB 1 SHQESTYGRSRGSGSGS 19
 RESULT 13
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.
 XX
 AC AAW61513;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc9, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 7
 FT Modified-site /note= "Citruilline"
 FT Modified-site 16
 FT /note= "Citruilline"
 XX
 W09822503-A2.
 XX
 PN 28-MAY-1998.
 XX
 PD 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 DB 1 SHQSTXGRSGRSGSGS 19

RESULT 14
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

AC AAW61510;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide c1c6, based on cDNA of a proflaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"
 FT
 XX

XX W09822503-A2.
 XX
 XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 XX

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 DB 1 SHQSTXGRSGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide XI based on cDNA of a proflaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT
 XX

XX W09822503-A2.
 XX
 XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 XX

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC

XX Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 2.6e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGSGS 19
 DB 4 HOESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:18
 Job time : 54.1302 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds

(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88

Sequence: 1 SHOESTXGRSGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 segs, 321558718 residues 1349238

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-4
4	84	95.5	19	9	US-09-308-150-8
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	84	95.5	19	9	US-09-308-150-2
9	84	95.5	19	9	US-09-308-150-3
10	84	95.5	19	9	US-09-308-150-5
11	84	95.5	19	9	US-09-308-150-6
12	84	95.5	19	9	US-09-308-150-7
13	84	95.5	19	9	US-09-308-150-9
14	84	95.5	19	9	US-09-308-150-10
15	84	95.5	19	9	US-09-308-150-22

16	50	56.8	477	15	US-10-161-927-62	Sequence 62, Appl
17	47	53.4	436	15	US-10-369-493-8178	Sequence 8178, Ap
18	46	52.3	123	16	US-10-767-701-32436	Sequence 32436, A
19	45	51.1	78	16	US-10-437-963-203073	Sequence 203073,
20	45	51.1	105	16	US-10-437-963-196208	Sequence 196208,
21	44	50.0	133	16	US-10-437-963-136367	Sequence 136367,
22	44	50.0	1936	14	US-10-156-761-11212	Sequence 11212, A
23	44	50.0	1087	9	US-09-918-908-24	Sequence 24, Appl
24	44	50.0	1887	16	US-10-641-991-24	Sequence 24, Appl
25	43.5	49.4	92	16	US-10-437-963-118183	Sequence 118183,
26	43	48.9	103	12	US-10-424-559-257534	Sequence 257534,
27	43	48.9	185	16	US-10-437-963-137528	Sequence 137528,
28	43	48.9	185	16	US-10-437-963-137528	Sequence 137528,
29	43	48.9	570	10	US-10-408-765A-2385	Sequence 2385, Ap
30	43	48.9	694	14	US-09-847-102A-43	Sequence 43, Appl
31	43	48.9	797	14	US-10-152-5148-2	Sequence 2, Appl1
32	43	48.9	889	16	US-10-032-585-7208	Sequence 7208, Ap
33	42	47.7	121	12	US-10-437-963-181020	Sequence 181020,
34	42	47.7	156	16	US-10-424-559-203241	Sequence 203241,
35	42	47.7	176	12	US-10-437-963-200525	Sequence 200525,
36	42	47.7	176	12	US-10-287-122A-76380	Sequence 76380, A
37	42	47.7	179	12	US-10-425-114-53545	Sequence 53545, A
38	42	47.7	588	16	US-10-437-963-185974	Sequence 185974,
39	42	47.7	618	16	US-10-437-963-177280	Sequence 177280,
40	42	47.7	772	15	US-10-767-701-45284	Sequence 45284, A
41	41.5	47.2	136	16	US-10-369-493-6474	Sequence 6474, Ap
42	41.5	47.2	19723	15	US-10-767-701-40697	Sequence 40697, A
43	41	46.6	50	14	US-10-084-846A-5	Sequence 5, Appl1
44	41	46.6	178	12	US-10-029-386-27692	Sequence 27692, A
45	41	46.6	184	9	US-10-425-114-69976	Sequence 69976, A
					US-09-925-300-1619	Sequence 1619, Ap

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known CDNA sequences of human profilaggrin
US-09-308-150-12
Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3.9e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 SHOESTXGRSGRSGSGS 19

Db 1 SH0ESTAGRSRGRSGSGS 19

RESULT 2

US-09-308-150-1
Sequence 1, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: US/09/308,150
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 5.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTYGRSRGXSGRSGS 19
Db 1 SH0ESTYGRSRGRSGSGS 19

RESULT 3

US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: US/09/308,150
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 5.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTYGRSRGXSGRSGS 19
Db 1 SH0ESTYGRSRGRSGSGS 19

RESULT 4

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: US/09/308,150
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-8

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SH0ESTYGRSRGXSGRSGS 19
Db 1 SH0ESTYGRSRGXSGRSGS 19

RESULT 5

US-09-308-150-11
Sequence 11, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: US/09/308,150
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 19

;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; PRIOR FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-11

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 6
US-09-308-150-13
;; Sequence 13, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: known CDNA sequences of human profilaggrin
US-09-308-150-13

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 7
US-09-308-150-14

;; Sequence 14, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 14
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: known CDNA sequences of human profilaggrin
US-09-308-150-14

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 8
US-09-308-150-2
;; Sequence 2, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: known CDNA sequences of human profilaggrin
;; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 9

US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 10

US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 11

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 12

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 SHOESTXGRSRGSGSGS 19
Db 1 SHOESTXGRSRGSGSGS 19

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Cy 1 SHOESTXGRSRGSGSGS 19
Db 1 SHOESTXGRSRGSGSGS 19

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3) .. (16)
US-09-308-150-10

Query Match      85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 HOESTXGRSRGSGSGS 19
Db 4 HOESTXGRSRGSGSGS 21

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NUS00 INNS:031
; CURRENT APPLICATION NUMBER: US/09/747,029A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```

/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic Peptides
/ NAME/KEY: MOD_RES
/ LOCATION: (10)..(10)
/ OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

Query Match 85.2%; Score 75; DB 9; Length 22;
Best Local Similarity 88.9%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HQESTXGRSRGSGSGS 19
Db 5 HQESTXGRSRGSGSGS 22

Search completed: September 28, 2004, 07:28:52
Job time : 112.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67,892 Million cell updates/sec

Title: US-09-308-150-8
Perfect score: 88
Sequence: 1 SH0ESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgnt2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgnt2_6/prodata/2/1aa/5A_COMB.pep:*
3: /cgnt2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgnt2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgnt2_6/prodata/2/1aa/5B_COMB.pep:*
6: /cgnt2_6/prodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	52.3	491	US-09-489-039A-7836	Sequence 7836, App
2	44	50.0	546	US-09-252-991A-24620	Sequence 24620, A
3	42	47.7	395	US-09-252-991A-21702	Sequence 21702, A
4	41	46.6	237	US-09-252-991A-23304	Sequence 23304, A
5	41	46.6	255	US-09-252-991A-19598	Sequence 19598, A
6	41	46.6	421	US-09-252-991A-32326	Sequence 32326, A
7	41	46.6	843	US-08-591-502B-45	Sequence 45, Appl
8	40	45.5	213	US-09-252-991A-32093	Sequence 32093, A
9	40	45.5	465	US-09-252-991A-18919	Sequence 18919, A
10	40	45.5	480	US-09-252-991A-22879	Sequence 22879, A
11	40	45.5	486	US-08-821-355A-8	Sequence 8, Appl
12	40	45.5	486	US-09-003-687A-8	Sequence 8, Appl
13	40	45.5	486	US-09-136-605-8	Sequence 8, Appl
14	40	45.5	660	US-09-058-489-18	Sequence 18, Appl
15	40	45.5	660	US-09-058-489-91	Sequence 91, Appl
16	40	45.5	660	US-09-976-594-787	Sequence 787, App
17	40	45.5	730	US-08-591-502B-49	Sequence 49, Appl
18	40	45.5	842	US-08-591-502B-50	Sequence 50, Appl
19	40	45.5	842	US-08-591-502B-51	Sequence 51, Appl
20	40	45.5	843	US-09-119-528A-2	Sequence 2, Appl
21	40	45.5	843	US-08-591-502B-48	Sequence 48, Appl
22	40	45.5	843	US-08-591-502B-59	Sequence 59, Appl
23	40	45.5	845	US-08-591-502B-46	Sequence 46, Appl
24	40	45.5	1293	US-09-170-496D-292	Sequence 292, App
25	40	45.5	1293	US-09-364-425B-57	Sequence 57, Appl
26	39.5	44.9	191	US-09-252-991A-18626	Sequence 18626, A
27	39	44.3	250	US-09-252-991A-31343	Sequence 31343, A

28	39	44.3	476	US-09-252-991A-18968	Sequence 18968, A
29	39	44.3	550	US-09-252-991A-25062	Sequence 25062, A
30	38	43.2	116	US-09-489-039A-13128	Sequence 13128, A
31	38	43.2	146	US-09-252-991A-21166	Sequence 21166, A
32	38	43.2	175	US-09-252-991A-19711	Sequence 19711, A
33	38	43.2	215	US-09-252-991A-30077	Sequence 30077, A
34	38	43.2	329	US-09-252-991A-28988	Sequence 28988, A
35	38	43.2	339	US-09-252-991A-21715	Sequence 21715, A
36	38	43.2	359	US-09-252-991A-20245	Sequence 20245, A
37	38	43.2	430	US-09-252-991A-22599	Sequence 22599, A
38	38	43.2	508	US-08-818-024-3	Sequence 3, Appl
39	38	43.2	508	US-09-334-775A-3	Sequence 3, Appl
40	38	43.2	518	US-09-252-991A-25967	Sequence 25967, A
41	38	43.2	580	US-08-482-677-10	Sequence 10, Appl
42	38	43.2	580	US-10-033-174-10	Sequence 10, Appl
43	38	43.2	638	US-09-252-991A-18036	Sequence 18036, A
44	38	43.2	735	US-09-252-991A-18657	Sequence 18657, A
45	38	43.2	774	US-09-252-991A-16789	Sequence 16789, A

ALIGNMENTS

```

RESULT 1
US-09-489-039A-7836
Sequence 7836, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,20040001
CURRENT FILING DATE: 2000-01-27
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7836
LENGTH: 491
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7836

Query Match      52.3%, Score 46, DB 4, Length 491,
Best Local Similarity 69.2%, Pred. No. 10,
Matches 9, Conservative 1, Mismatches 3, Indels 0, Gaps 0,

Cy      5 STXGRSGRSGRS 17
Db      217 TASGRSGRSGRS 229

RESULT 2
US-09-252-991A-24620
Sequence 24620, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24620
LENGTH: 546
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24620

```

Query Match 50.0%; Score 44; DB 4; Length 546;
 Best Local Similarity 64.3%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSG 18
 |||:|||||
 Db 427 SRLGRDRGASGRPG 440

RESULT 3
 US-09-252-991A-21702
 ; Sequence 21702, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-07-27
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21702
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21702

Query Match 47.7%; Score 42; DB 4; Length 395;
 Best Local Similarity 50.0%; Pred. No. 36;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGXSGRSG 18
 |||:|||||
 Db 251 SHQASERGDHESGRPG 268

RESULT 4
 US-09-252-991A-23304
 ; Sequence 23304, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 23304
 ; LENGTH: 237
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-23304

Query Match 46.6%; Score 41; DB 4; Length 237;
 Best Local Similarity 47.1%; Pred. No. 31;
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 HSESTXGRSGXSGRSG 18
 |||:|||||
 Db 124 HQRHRRGROGAGADPG 140

RESULT 5

US-09-252-991A-19598
 ; Sequence 19598, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 19598
 ; LENGTH: 255
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-19598

Query Match 46.6%; Score 41; DB 4; Length 255;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 GSRSGXSGRSGS 19
 |||:|||||
 Db 64 GDRGRAGRAGA 75

RESULT 6
 US-09-252-991A-32326
 ; Sequence 32326, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32326
 ; LENGTH: 421
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-32326

Query Match 46.6%; Score 41; DB 4; Length 421;
 Best Local Similarity 47.1%; Pred. No. 55;
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 HSESTXGRSGXSGRSG 18
 |||:|||||
 Db 13 HARTGTGRRGRGRSG 29

RESULT 7
 US-08-591-502B-45
 ; Sequence 45, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytopathic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor

```

; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: NO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauber
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45
;
; Query Match
; Best Local Similarity 46.6%; Score 41; DB 4; Length 843;
; Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
;
; QY 3 QESTXGRSGXSGRSGS 19
; DB 226 QPQGGMARGSGRSGS 242
;
; RESULT 8
; US-09-252-991A-32093
; Sequence 32093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32093
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-32093
;
; Query Match
; Best Local Similarity 45.5%; Score 40; DB 4; Length 213;
; Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

; QY 3 QESTXGRSGXSGRSGS 19
; DB 182 RETTPRGDGGGRGT 198
;
; RESULT 9
; US-09-252-991A-18919
; Sequence 18919, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18919
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18919
;
; Query Match
; Best Local Similarity 45.5%; Score 40; DB 4; Length 465;
; Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
;
; QY 1 SHQSTXGRSGXSGRSGS 19
; DB 83 SCGTGCRGALRGSGRAGS 101
;
; RESULT 10
; US-09-252-991A-22879
; Sequence 22879, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22879
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22879
;
; Query Match
; Best Local Similarity 46.7%; Score 40; DB 4; Length 480;
; Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
;
; QY 1 SHQSTXGRSGXSG 15
; DB 52 SKDSSAGRGARSTG 66
;
; RESULT 11
; US-08-821-355A-8
; Sequence 8, Application US/08821355A
; Patent No. 5851775
; GENERAL INFORMATION:

```

APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821.355A
FILING DATE: 20-MAR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107.05064
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5851775e
US-08-821-355A-8
Query Match 45.5%; Score 40; DB 2; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CY 2 HOESTXGRSGXSG 15
DB 383 HOESTTGKRNFAFG 396
RESULT 12
US-09-003-687A-8
Sequence 8, Application US/09003687A
Patent No. 5998600
GENERAL INFORMATION:
APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003.687A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/821.355
FILING DATE: 20-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107.05064
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998600e
US-09-003-687A-8
Query Match 45.5%; Score 40; DB 2; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CY 2 HOESTXGRSGXSG 15
DB 383 HOESTTGKRNFAFG 396
RESULT 13
US-09-136-605-8
Sequence 8, Application US/09136605A
Patent No. 6140052
GENERAL INFORMATION:
APPLICANT: He, Tong-Chuan
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
FILE REFERENCE: 1107.75741
CURRENT APPLICATION NUMBER: US/09/136.605A
CURRENT FILING DATE: 1998-08-20
EARLIER APPLICATION NUMBER: 08/821.355
EARLIER FILING DATE: 1997-03-20
EARLIER APPLICATION NUMBER: 09/003.687
EARLIER FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 486
TYPE: PRT
ORGANISM: Homo sapiens
US-09-136-605-8
Query Match 45.5%; Score 40; DB 3; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CY 2 HOESTXGRSGXSG 15

Db 383 H0ESTTGKRNFG 396

```

RESULT 14
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lam, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 660
; TYPE: PRF
; ORGANISM: Human
US-09-058-489-18

```

Query Match 45.5%; Score 40; DB 3; Length 660;
 Best Local Similarity 64.3%; Pred. No. 1.2e+02;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSG 18
 Db 609 SGFGASRGSSSRSG 622

```

RESULT 15
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lam, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 660
; TYPE: PRF
; ORGANISM: Human
US-09-058-489-91

```

Query Match 45.5%; Score 40; DB 3; Length 660;
 Best Local Similarity 64.3%; Pred. No. 1.2e+02;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSG 18
 Db 609 SGFGASRGSSSRSG 622

Search completed: September 28, 2004, 06:26:45
 Job time : 14.4479 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-9

Perfect score: 88

Sequence: 1 SHQESTXGSRGRSGXSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 45 summaries

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	90.9	2248	2	A35938
2	78	88.6	416	2	A32947
3	69	78.4	591	2	A45135
4	52	59.1	506	1	W2W47
5	49	55.7	822	2	T51049
6	44	50.0	745	2	G01025
7	44	50.0	774	2	I48609
8	44	50.0	1119	2	T16720
9	43	48.9	306	2	T21220
10	43	48.9	694	2	S71786
11	43	48.9	843	2	H86209
12	43	48.9	849	2	A96592
13	43	48.9	1015	2	I42013
14	42	47.7	136	2	T35632
15	42	47.7	176	2	F71370
16	42	47.7	308	2	A29379
17	42	47.7	315	2	S53889
18	42	47.7	772	2	T27512
19	42	47.7	2233	2	S63347
20	42	47.7	2420	2	A84852
21	41	46.6	167	2	T52599
22	41	46.6	174	2	T52600
23	41	46.6	373	2	T02976
24	41	46.6	471	2	T33997
25	41	46.6	528	2	G02127
26	41	46.6	759	2	P86362
27	41	46.6	825	2	UC4163
28	41	46.6	836	2	G84727
29	41	46.6	880	2	T04523

30	40.5	46.0	625	2	A34615	profilaggrin - rat
31	40	45.5	259	2	C69122	hypothetical prote
32	40	45.5	312	2	A31846	130K paracrystall
33	40	45.5	331	2	F63142	F631.5 protein - A
34	40	45.5	346	2	A82623	Tola protein Xf189
35	40	45.5	399	2	T46259	hypothetical prote
36	40	45.5	462	4	S33798	FMS/CHOP mutant fu
37	40	45.5	467	2	T01462	hypothetical prote
38	40	45.5	526	1	S33799	RNA-binding protei
39	40	45.5	526	2	A41698	cell division cont
40	40	45.5	807	2	T02738	probable AAA-type
41	40	45.5	854	2	S61164	repeat/recombinati
42	40	45.5	1557	2	T02859	probable serine/th
43	38.5	44.9	135	2	A46398	RNA-binding protei
44	38.5	44.9	286	2	S50855	neurotrophin-6 - s
45	39.5	44.9	1895	2	T06609	disease resistance

ALIGNMENTS

RESULT 1

A35938

profilaggrin - human (fragments)

C:Species: Homo sapiens (man)

C>Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

C:Accession: A35938

R:Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.

Biochemistry 29, 9432-9440, 1990

A>Title: Organization, structure, and polymorphisms of the human profilaggrin gene.

A:Reference number: A35938; PMID:91064347; PMID:2248957

A:Accession: A35938

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2248 <GNN>

A:Cross-references: GB:U02929

C:Genetics:

A:Gene: GDB:FLG

A:Cross-references: GDB:119912; OMIM:135940

A:Map position: 1q21-1q21

C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

F:246-569/Region: profilaggrin repeat

F:570-893/Region: profilaggrin repeat

F:1074-1397/Region: profilaggrin repeat

F:1573-1896/Region: profilaggrin repeat

Query Match 90.9%; Score 80; DB 2; Length 2248;
Best Local Similarity 84.2%; Pred. No. 0.00011;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQESTXGSRGRSGXSGS 19
Db 551 SHQESTXGSRGRSGXSGS 569

RESULT 2

A32947

profilaggrin precursor - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999

C:Accession: A32947

R:McKinley-Grant, L.U.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc

Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989

A>Title: Characterization of a cDNA clone encoding human profilaggrin and localization of th

A:Reference number: A32947; PMID:89296901; PMID:2740331

A:Accession: A32947

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <MKC>

A:Cross-references: GB:M24355; NID:G182604; PIDN:AAA52454.1; PID:G182605

A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue

C:Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 88.6%; Score 78; DB 2; Length 416;
 Best Local Similarity 84.2%; Pred. No. 4.7e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 DB 7 SHQSTXGRSGRSGSGS 25

RESULT 3

A45135
 C:Species: Homo sapiens (man)
 C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135
 R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nirunskiri, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992

A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717

A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101089; GB:M90697; NID:9190408; PID:AAA60177.1; PID:9553621
 A:Note: sequence extracted from NCBI backbone (NCBIF:118773)

C:Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940

A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 78.4%; Score 69; DB 2; Length 591;
 Best Local Similarity 78.9%; Pred. No. 0.0019;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 DB 449 SHQSTXGRSGRSGSGS 467

RESULT 4

M2ML47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D53324
 R:Kiyono, T.; Adachi, A.; Ishibashi, K.
 Virology 177, 401-405, 1990

A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
 A:Reference number: A35324; MUID:90281611; PMID:2162112

A:Accession: D53324
 A:Status: translation not shown

A:Molecule type: DNA
 A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M2305; NID:9333062; PID:AAA6979.1; PID:9333067
 C:Superfamily: Papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 59.1%; Score 52; DB 1; Length 506;
 Best Local Similarity 52.6%; Pred. No. 0.83;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19

DB 342 SREGTRGRGRGRGRGS 360

RESULT 5

T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12FL10
 C:Species: Neurospora crassa
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049
 R:Schulte, U.; Aign, V.; Hehse, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000

A:Reference number: Z55286
 A:Accession: T51049

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AJ390091; GSPDB:GN00116; NCSP:B12FL10

A:Experimental source: BAC clone B12FL; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B12FL10
 A:Map position: 6
 A:Introns: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 4.3;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGSGS 19
 DB 414 RESAGRTGRGRGRGT 430

RESULT 6

G01025
 Serine/threonine protein kinase - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998
 C:Accession: G01025
 R:Navarro, E.
 submitted to the EMBL Data Library, April 1996

A:Reference number: H00564
 A:Accession: G01025
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-745 <NAV>
 A:Cross-references: EMBL:X97630; NID:91310674

A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 F:18-271/Domain: protein kinase homology <KIN>

Query Match 50.0%; Score 44; DB 2; Length 745;
 Best Local Similarity 60.0%; Pred. No. 25;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGSGS 19
 DB 585 SPSCGSGRGRGSGS 599

RESULT 7

I48609
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
 C:Accession: I48609; S91333
 R:Ingilis, J.D.; Lee, M.; Hill, R.E.
 Mamm. Genome 4, 401-403, 1993

A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
 A:Reference number: I48609; MUID:93364122; PMID:8358177

A:Accession: I48609
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-774 <INGI>

A/Cross-references: EMBL:X70764; NID:g57919; PIDN:CAA50040.1; PID:g57920
 R:Jinglis, J.D.; Lee, M.; Hill, R.E.
 Submitted to the EMBL Data Library, January 1993
 A/Description: A novel protein kinase with homologues in yeast maps to mouse chromosome
 A/Reference number: S31333
 A/Accession: S31333
 A/Molecule type: mRNA
 A/Residues: 1-698, 'X', '700-702', 'GRLPLPAKASCTSGNVCAGMAHQATRR', '731-774' <IMG2>
 A/Cross-references: EMBL:X70764
 C/Genetics:
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 C/51-304/Domain: protein kinase homology <KIN>
 F:59-67/Region: protein kinase ATP-binding motif

Query Match 50.0%; Score 44; DB 2; Length 774;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 5 STXGRSGRSGXSGS 19
 DB 614 SPGCHGGRGASGS 628

RESULT 8
 T16720
 hypothetical protein R11G1.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C/Accession: T16720
 R:Miller, N.
 Submitted to the EMBL Data Library, November 1995
 A/Description: The sequence of C. elegans cosmid R11G1.
 A/Reference number: Z18566
 A/Accession: T16720
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1119 <MTL>
 A/Cross-references: EMBL:U41016; NID:g1086652; PID:g1086656; PIDN:AAA82321.1; CESP:R11G1
 C/Genetics:
 A/Gene: CESP:R11G1.1
 A/Introns: 24/1; 65/1; 107/1; 145/1; 500/1; 661/1; 693/3; 746/1; 802/2; 857/3; 926/3; 97

Query Match 50.0%; Score 44; DB 2; Length 1119;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 4 ESTXGRSGRSGXSGS 19
 DB 1008 EETTKENRREGGGS 1023

RESULT 9
 T21220
 hypothetical protein F21H7.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 R:Gardner, A.
 Submitted to the EMBL Data Library, March 1997
 A/Reference number: Z19393
 A/Accession: T21220
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-306 <MTL>
 A/Cross-references: EMBL:Z93379; PIDN:CA907586.1; GSPDB:GN00023; CESP:F21H7.5
 A/Experimental source: clone F21H7
 C/Genetics:
 A/Gene: CESP:F21H7.5
 A/Map position: 5
 A/Introns: 28/3; 146/3; 232/3
 C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGXSG 18
 DB 48 SHRSKTKRKSGKSGSG 65

RESULT 10
 S71786
 wingless receptor precursor dfz2 - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 R:Rahmou, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang, Y.; Macke, J.P.; Andrew,
 Nature 382, 225-230, 1996
 A/Title: A new member of the frizzled family from Drosophila functions as a wingless rec
 A/Reference number: S71786; MUID:96353971; PMID:8717036
 A/Accession: S71786
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-694 <BHA>
 A/Cross-references: EMBL:U65589
 A/Note: mRNA was also sequenced
 R:Bhanot, P.; Wang, Y.; Nathans, J.
 Submitted to the EMBL Data Library, July 1996
 A/Reference number: S78444
 A/Accession: S78444
 A/Molecule type: DNA
 A/Residues: 1-416, 'T', '418-694' <BHW>
 A/Cross-references: EMBL:U65589; NID:g1518050; PIDN:AA047273.1; PID:g1518051
 C/Genetics:
 A/Gene: dfz2
 A/Cross-references: Flybase:FBgn0016797
 C/Superfamily: fruit fly frizzled protein
 C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGXSGS 19
 DB 182 SYTEAGSGSGSGSGSGSGS 200

RESULT 11
 H86209
 protein F22G5.10 (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: H86209
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 anser, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: H86209
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-843 <STO>
 A/Cross-references: GB:AE005172; NID:g8778569; PIDN:AAF79577.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F22G5.10
 A/Map position: 1

Query March	48.9%;	Score 43;	DB 2;	Length 843;
Best Local Similarity	72.7%;	Pred. No. 41;		
Matches	8;	Conservative	1;	Mismatches 2;
				Indels 0;
				Gaps 0;
Q7	8	GRAGRGSGXG	18	
Db	440	GRAGRGGFSG	450	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-9

Sequence: 1 SHQESTXGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	88.6	416	1	PI1A_HUMAN
2	52	59.1	506	1	VE2_HPV47
3	44	50.0	774	1	MRK2_MOUSE
4	43	48.9	694	1	FRZ2_DROME
5	43	48.9	1015	1	FRQ_CREP
6	42	47.7	176	1	SSB_TREPA
7	42	47.7	308	1	GRP3_ARTSA
8	42	47.7	315	1	SOL2_YEAST
9	42	47.7	772	1	MR11_CAEEL
10	42	47.7	2233	1	COAC_YEAST
11	41	46.6	825	1	SE5_RAT
12	40.5	46.0	379	1	ROA3_MOUSE
13	40	45.5	255	1	THO4_MOUSE
14	40	45.5	257	1	THO4_HUMAN
15	40	45.5	380	1	PEXD_PICPA
16	40	45.5	526	1	ZABR_YEAST
17	40	45.5	526	1	FUS_HUMAN
18	40	45.5	854	1	XRS2_YEAST
19	39.5	44.9	135	1	RBP1_DROME
20	39.5	44.9	1895	1	WR19_ARATH
21	39.5	44.9	2594	1	7LES_DROVT
22	39	44.3	70	1	L2WU_ADE40
23	39	44.3	271	1	HEXP_LEIMA
24	39	44.3	419	1	TCF7_MOUSE
25	39	44.3	617	1	DBP1_YEAST
26	39	44.3	686	1	BDP1_YEAST
27	39	44.3	978	1	PEX6_RAT
28	39	44.3	980	1	PEX6_HUMAN
29	39	44.3	1703	1	SNF2_YEAST
30	38.5	43.8	697	1	AN3_XENLA
31	38	43.2	168	1	DH1_MAIZE
32	38	43.2	452	1	VE2_HPV17
33	38	43.2	483	1	VE2_HPV14

Result	ID	Score	Query Match	Length	DB ID	Description
34	38	43.2	514	1	VE2_HPV05	P06921 human papill
35	38	43.2	614	1	DDX5_HUMAN	P17844 homo sapien
36	38	43.2	614	1	DDX5_MOUSE	O61656 mus musculu
37	38	43.2	734	1	ZN42_HUMAN	P28698 homo sapien
38	38	43.2	775	1	ICP0_HSV1	P08393 herpes simp
39	38	43.2	843	1	DPOL_HPV4	P03157 hepatitis b
40	38	43.2	955	1	CLU2_HUMAN	O29440 homo sapien
41	38	43.2	1122	1	DSG2_MOUSE	O55111 mus musculu
42	38	43.2	1232	1	YOO5_CAEEL	P34643 caenorhabdi
43	38	43.2	1337	1	DEXT_STRO	P39653 streptococ
44	38	43.2	3579	1	STAN_DROME	O95588 drosophila
45	37	42.0	123	1	LSM4_CAEEL	Q19552 caenorhabdi

ALIGNMENTS

Result	ID	Score	Query Match	Length	DB ID	Description
1	PI1A_HUMAN	78	88.6	416	1	PI1A_HUMAN
2	VE2_HPV47	52	59.1	506	1	VE2_HPV47
3	MRK2_MOUSE	44	50.0	774	1	MRK2_MOUSE
4	FRZ2_DROME	43	48.9	694	1	FRZ2_DROME
5	FRQ_CREP	43	48.9	1015	1	FRQ_CREP
6	SSB_TREPA	42	47.7	176	1	SSB_TREPA
7	GRP3_ARTSA	42	47.7	308	1	GRP3_ARTSA
8	SOL2_YEAST	42	47.7	315	1	SOL2_YEAST
9	MR11_CAEEL	42	47.7	772	1	MR11_CAEEL
10	COAC_YEAST	42	47.7	2233	1	COAC_YEAST
11	SE5_RAT	41	46.6	825	1	SE5_RAT
12	ROA3_MOUSE	40.5	46.0	379	1	ROA3_MOUSE
13	THO4_MOUSE	40	45.5	255	1	THO4_MOUSE
14	THO4_HUMAN	40	45.5	257	1	THO4_HUMAN
15	PEXD_PICPA	40	45.5	380	1	PEXD_PICPA
16	ZABR_YEAST	40	45.5	526	1	ZABR_YEAST
17	FUS_HUMAN	40	45.5	526	1	FUS_HUMAN
18	XRS2_YEAST	40	45.5	854	1	XRS2_YEAST
19	RBP1_DROME	39.5	44.9	135	1	RBP1_DROME
20	WR19_ARATH	39.5	44.9	1895	1	WR19_ARATH
21	7LES_DROVT	39.5	44.9	2594	1	7LES_DROVT
22	L2WU_ADE40	39	44.3	70	1	L2WU_ADE40
23	HEXP_LEIMA	39	44.3	271	1	HEXP_LEIMA
24	TCF7_MOUSE	39	44.3	419	1	TCF7_MOUSE
25	DBP1_YEAST	39	44.3	617	1	DBP1_YEAST
26	BDP1_YEAST	39	44.3	686	1	BDP1_YEAST
27	PEX6_RAT	39	44.3	978	1	PEX6_RAT
28	PEX6_HUMAN	39	44.3	980	1	PEX6_HUMAN
29	SNF2_YEAST	39	44.3	1703	1	SNF2_YEAST
30	AN3_XENLA	38.5	43.8	697	1	AN3_XENLA
31	DH1_MAIZE	38	43.2	168	1	DH1_MAIZE
32	VE2_HPV17	38	43.2	452	1	VE2_HPV17
33	VE2_HPV14	38	43.2	483	1	VE2_HPV14

DR PFam; PF03516; Filaggrin; 2.
DR PRINTS; PR00487; FILAGRIN.
KW Phosphorylation; Citrullination; Developmental protein.
FT NON TER 1
SQ SEQUENCE 416 AA; 44105 MW; DERA3218BA043F32 CRC64;

Query Match 88.6%; Score 76; DB 1; Length 416;
Best Local Similarity 84.2%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGXSGS 19
7 SHQSTXGRSGRSGRSGS 25

RESULT 2
VE2_HPVA7 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
RT type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCCNNNNNGT-3'), PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -.
DR FIR; D35324; W2ML47.
DR HSSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR003021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C; 1.
DR Pfam; PF00508; E2_N; 1.
DR ProDom; PD000672; E2_C; 1.
DR ProDom; PD000678; E2_N; 1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BB725065 CRC64;

Query Match 59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.33;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGXSGS 19

DB 342 SREGNTRGRGRGROGRAGS 360

RESULT 3
MRK2_MOUSE STANDARD; PRT; 774 AA.
ID MRK2_MOUSE
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
GN MRK2 OR EMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Embryo;
RX MEDLINE=93364122; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19.";
RL Mamm. Genome 4:401-403(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70764; CA50040.1; -.
DR FIR; 148609; 148609.
DR HSSP; Q63450; 1A06.
DR MGD; MGI:99638; Mark2.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 53 304
FT NP_BIND 59 67
FT BINDING 82 82
FT ACT_SITE 175 175
FT ACT_SITE BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BR8DBF44343A CRC64;

Query Match 50.0%; Score 44; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGRSGXSGS 19
614 SPFGHSGRGRAGS 628

RESULT 4
FRZ2_DROME

ID FR22_DROME STANDARD; PRT; 694 AA.
 AC Q9VAX3; Q94916; Q9VAX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frizzled protein 2 precursor (Frizzled-2) (dfrz2).
 GN FZ2 OR CG9739/CG14083.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhanot P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.;
 RT "A new member of the frizzled family from Drosophila functions as a
 RT Wingless receptor";
 RL Nature 362:225-230(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceolnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.,
 RA Sutton G.C., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.G., Rogers Y.-H.C., Blaise J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Arlitt J.F., Agbayani A., An H.-J., Andrews-Franko C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borokova D., Buchan M.R., Buck J., Brokstein P., Brothier P.,
 RA Chertis J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Curry J.M., Cawley S., Dahlke C., Davernport L.B., Davies P.,
 RA de Pablos G., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hosten D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemsinn J.A., Kechum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhina N.V., Modyarty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins, inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Recruited to coordinate
 CC the cytoskeletons of epidermal cells to produce a parallel array

CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage
 CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut and primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U65589; AAC47273.1; -
 CC EMBL; AB003518; AAF49185.2; -
 CC F1R; S71786; S71786.
 CC FLYBase; FBgn0016797; f22.
 CC GO; GO:0016021; C:integral to membrane; NNS.
 CC GO; GO:0017147; F:Wnt-protein binding; IDA.
 CC GO; GO:0006928; P:cell motility; IMP.
 CC GO; GO:0007163; P:establishment and/or maintenance of cell po. ; ISS.
 CC GO; GO:0008585; P:female gonad development; IMP.
 CC GO; GO:0016055; P:Wnt receptor signaling pathway; IDA.
 CC InterPro; IPR000539; Frizzled.
 CC InterPro; IPR000024; Fz_domain.
 CC InterPro; IPR000832; GPCR_secretin.
 CC Pfam; PF01534; Frizzled_1.
 CC Pfam; PF01392; Fz_1.
 CC PRINTS; PRO0489; FRIZZLED.
 CC SMART; SM00063; FRI; 1.
 CC PROSITE; PSS0038; FZ; 1.
 CC PROSITE; PSS0261; G_PROTEIN_RECEP_F2_4; 1.
 CC KW MultiGene Family; Receptor; G-protein coupled receptor; Transmembrane;
 CC Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 CC FT SIGNAL 1 22
 CC FT CHAIN 23 694
 CC FT DOMAIN 23 315
 CC FT TRANSMEM 316 336
 CC FT DOMAIN 337 352
 CC FT TRANSMEM 353 373
 CC FT DOMAIN 374 397
 CC FT TRANSMEM 398 418
 CC FT DOMAIN 419 439
 CC FT TRANSMEM 440 460
 CC FT DOMAIN 461 482
 CC FT TRANSMEM 483 503
 CC FT DOMAIN 504 534
 CC FT TRANSMEM 535 555
 CC FT DOMAIN 556 584
 CC FT TRANSMEM 585 605
 CC FT DOMAIN 606 694
 CC FT TRANSMEM 187 225
 CC FT DOMAIN 187 225
 CC FT TRANSMEM 226 246
 CC FT DOMAIN 247 267
 CC FT TRANSMEM 268 288
 CC FT DOMAIN 289 309
 CC FT TRANSMEM 310 330
 CC FT DOMAIN 331 351
 CC FT TRANSMEM 352 372
 CC FT DOMAIN 373 393
 CC FT TRANSMEM 394 414
 CC FT DOMAIN 415 435
 CC FT TRANSMEM 436 456
 CC FT DOMAIN 457 477
 CC FT TRANSMEM 478 498
 CC FT DOMAIN 499 519
 CC FT TRANSMEM 520 540
 CC FT DOMAIN 541 561
 CC FT TRANSMEM 562 582
 CC FT DOMAIN 583 603
 CC FT TRANSMEM 604 624
 CC FT DOMAIN 625 645
 CC FT TRANSMEM 646 666
 CC FT DOMAIN 667 687
 CC FT TRANSMEM 688 708
 CC FT DOMAIN 709 729
 CC FT TRANSMEM 730 750
 CC FT DOMAIN 751 771
 CC FT TRANSMEM 772 792
 CC FT DOMAIN 793 813
 CC FT TRANSMEM 814 834
 CC FT DOMAIN 835 855
 CC FT TRANSMEM 856 876
 CC FT DOMAIN 877 897
 CC FT TRANSMEM 898 918
 CC FT DOMAIN 919 939
 CC FT TRANSMEM 940 960
 CC FT DOMAIN 961 981
 CC FT TRANSMEM 982 1002
 CC FT DOMAIN 1003 1023
 CC FT TRANSMEM 1024 1044
 CC FT DOMAIN 1045 1065
 CC FT TRANSMEM 1066 1086
 CC FT DOMAIN 1087 1107
 CC FT TRANSMEM 1108 1128
 CC FT DOMAIN 1129 1149
 CC FT TRANSMEM 1150 1170
 CC FT DOMAIN 1171 1191
 CC FT TRANSMEM 1192 1212
 CC FT DOMAIN 1213 1233
 CC FT TRANSMEM 1234 1254
 CC FT DOMAIN 1255 1275
 CC FT TRANSMEM 1276 1296
 CC FT DOMAIN 1297 1317
 CC FT TRANSMEM 1318 1338
 CC FT DOMAIN 1339 1359
 CC FT TRANSMEM 1360 1380
 CC FT DOMAIN 1381 1401
 CC FT TRANSMEM 1402 1422
 CC FT DOMAIN 1423 1443
 CC FT TRANSMEM 1444 1464
 CC FT DOMAIN 1465 1485
 CC FT TRANSMEM 1486 1506
 CC FT DOMAIN 1507 1527
 CC FT TRANSMEM 1528 1548
 CC FT DOMAIN 1549 1569
 CC FT TRANSMEM 1570 1590
 CC FT DOMAIN 1591 1611
 CC FT TRANSMEM 1612 1632
 CC FT DOMAIN 1633 1653
 CC FT TRANSMEM 1654 1674
 CC FT DOMAIN 1675 1695
 CC FT TRANSMEM 1696 1716
 CC FT DOMAIN 1717 1737
 CC FT TRANSMEM 1738 1758
 CC FT DOMAIN 1759 1779
 CC FT TRANSMEM 1780 1800
 CC FT DOMAIN 1801 1821
 CC FT TRANSMEM 1822 1842
 CC FT DOMAIN 1843 1863
 CC FT TRANSMEM 1864 1884
 CC FT DOMAIN 1885 1905
 CC FT TRANSMEM 1906 1926
 CC FT DOMAIN 1927 1947
 CC FT TRANSMEM 1948 1968
 CC FT DOMAIN 1969 1989
 CC FT TRANSMEM 1990 2010
 CC FT DOMAIN 2011 2031
 CC FT TRANSMEM 2032 2052
 CC FT DOMAIN 2053 2073
 CC FT TRANSMEM 2074 2094
 CC FT DOMAIN 2095 2115
 CC FT TRANSMEM 2116 2136
 CC FT DOMAIN 2137 2157
 CC FT TRANSMEM 2158 2178
 CC FT DOMAIN 2179 2199
 CC FT TRANSMEM 2200 2220
 CC FT DOMAIN 2221 2241
 CC FT TRANSMEM 2242 2262
 CC FT DOMAIN 2263 2283
 CC FT TRANSMEM 2284 2304
 CC FT DOMAIN 2305 2325
 CC FT TRANSMEM 2326 2346
 CC FT DOMAIN 2347 2367
 CC FT TRANSMEM 2368 2388
 CC FT DOMAIN 2389 2409
 CC FT TRANSMEM 2410 2430
 CC FT DOMAIN 2431 2451
 CC FT TRANSMEM 2452 2472
 CC FT DOMAIN 2473 2493
 CC FT TRANSMEM 2494 2514
 CC FT DOMAIN 2515 2535
 CC FT TRANSMEM 2536 2556
 CC FT DOMAIN 2557 2577
 CC FT TRANSMEM 2578 2598
 CC FT DOMAIN 2599 2619
 CC FT TRANSMEM 2620 2640
 CC FT DOMAIN 2641 2661
 CC FT TRANSMEM 2662 2682
 CC FT DOMAIN 2683 2703
 CC FT TRANSMEM 2704 2724
 CC FT DOMAIN 2725 2745
 CC FT TRANSMEM 2746 2766
 CC FT DOMAIN 2767 2787
 CC FT TRANSMEM 2788 2808
 CC FT DOMAIN 2809 2829
 CC FT TRANSMEM 2830 2850
 CC FT DOMAIN 2851 2871
 CC FT TRANSMEM 2872 2892
 CC FT DOMAIN 2893 2913
 CC FT TRANSMEM 2914 2934
 CC FT DOMAIN 2935 2955
 CC FT TRANSMEM 2956 2976
 CC FT DOMAIN 2977 2997
 CC FT TRANSMEM 2998 3018
 CC FT DOMAIN 3019 3039
 CC FT TRANSMEM 3040 3060
 CC FT DOMAIN 3061 3081
 CC FT TRANSMEM 3082 3102
 CC FT DOMAIN 3103 3123
 CC FT TRANSMEM 3124 3144
 CC FT DOMAIN 3145 3165
 CC FT TRANSMEM 3166 3186
 CC FT DOMAIN 3187 3207
 CC FT TRANSMEM 3208 3228
 CC FT DOMAIN 3229 3249
 CC FT TRANSMEM 3250 3270
 CC FT DOMAIN 3271 3291
 CC FT TRANSMEM 3292 3312
 CC FT DOMAIN 3313 3333
 CC FT TRANSMEM 3334 3354
 CC FT DOMAIN 3355 3375
 CC FT TRANSMEM 3376 3396
 CC FT DOMAIN 3397 3417
 CC FT TRANSMEM 3418 3438
 CC FT DOMAIN 3439 3459
 CC FT TRANSMEM 3460 3480
 CC FT DOMAIN 3481 3501
 CC FT TRANSMEM 3502 3522
 CC FT DOMAIN 3523 3543
 CC FT TRANSMEM 3544 3564
 CC FT DOMAIN 3565 3585
 CC FT TRANSMEM 3586 3606
 CC FT DOMAIN 3607 3627
 CC FT TRANSMEM 3628 3648
 CC FT DOMAIN 3649 3669
 CC FT TRANSMEM 3670 3690
 CC FT DOMAIN 3691 3711
 CC FT TRANSMEM 3712 3732
 CC FT DOMAIN 3733 3753
 CC FT TRANSMEM 3754 3774
 CC FT DOMAIN 3775 3795
 CC FT TRANSMEM 3796 3816
 CC FT DOMAIN 3817 3837
 CC FT TRANSMEM 3838 3858
 CC FT DOMAIN 3859 3879
 CC FT TRANSMEM 3880 3900
 CC FT DOMAIN 3901 3921
 CC FT TRANSMEM 3922 3942
 CC FT DOMAIN 3943 3963
 CC FT TRANSMEM 3964 3984
 CC FT DOMAIN 3985 4005
 CC FT TRANSMEM 4006 4026
 CC FT DOMAIN 4027 4047
 CC FT TRANSMEM 4048 4068
 CC FT DOMAIN 4069 4089
 CC FT TRANSMEM 4090 4110
 CC FT DOMAIN 4111 4131
 CC FT TRANSMEM 4132 4152
 CC FT DOMAIN 4153 4173
 CC FT TRANSMEM 4174 4194
 CC FT DOMAIN 4195 4215
 CC FT TRANSMEM 4216 4236
 CC FT DOMAIN 4237 4257
 CC FT TRANSMEM 4258 4278
 CC FT DOMAIN 4279 4299
 CC FT TRANSMEM 4300 4320
 CC FT DOMAIN 4321 4341
 CC FT TRANSMEM 4342 4362
 CC FT DOMAIN 4363 4383
 CC FT TRANSMEM 4384 4404
 CC FT DOMAIN 4405 4425
 CC FT TRANSMEM 4426 4446
 CC FT DOMAIN 4447 4467
 CC FT TRANSMEM 4468 4488
 CC FT DOMAIN 4489 4509
 CC FT TRANSMEM 4510 4530
 CC FT DOMAIN 4531 4551
 CC FT TRANSMEM 4552 4572
 CC FT DOMAIN 4573 4593
 CC FT TRANSMEM 4594 4614
 CC FT DOMAIN 4615 4635
 CC FT TRANSMEM 4636 4656
 CC FT DOMAIN 4657 4677
 CC FT TRANSMEM 4678 4698
 CC FT DOMAIN 4699 4719
 CC FT TRANSMEM 4720 4740
 CC FT DOMAIN 4741 4761
 CC FT TRANSMEM 4762 4782
 CC FT DOMAIN 4783 4803
 CC FT TRANSMEM 4804 4824
 CC FT DOMAIN 4825 4845
 CC FT TRANSMEM 4846 4866
 CC FT DOMAIN 4867 4887
 CC FT TRANSMEM 4888 4908
 CC FT DOMAIN 4909 4929
 CC FT TRANSMEM 4930 4950
 CC FT DOMAIN 4951 4971
 CC FT TRANSMEM 4972 4992
 CC FT DOMAIN 4993 5013
 CC FT TRANSMEM 5014 5034
 CC FT DOMAIN 5035 5055
 CC FT TRANSMEM 5056 5076
 CC FT DOMAIN 5077 5097
 CC FT TRANSMEM 5098 5118
 CC FT DOMAIN 5119 5139
 CC FT TRANSMEM 5140 5160
 CC FT DOMAIN 5161 5181
 CC FT TRANSMEM 5182 5202
 CC FT DOMAIN 5203 5223
 CC FT TRANSMEM 5224 5244
 CC FT DOMAIN 5245 5265
 CC FT TRANSMEM 5266 5286
 CC FT DOMAIN 5287 5307
 CC FT TRANSMEM 5308 5328
 CC FT DOMAIN 5329 5349
 CC FT TRANSMEM 5350 5370
 CC FT DOMAIN 5371 5391
 CC FT TRANSMEM 5392 5412
 CC FT DOMAIN 5413 5433
 CC FT TRANSMEM 5434 5454
 CC FT DOMAIN 5455 5475
 CC FT TRANSMEM 5476 5496
 CC FT DOMAIN 5497 5517
 CC FT TRANSMEM 5518 5538
 CC FT DOMAIN 5539 5559
 CC FT TRANSMEM 5560 5580
 CC FT DOMAIN 5581 5601
 CC FT TRANSMEM 5602 5622
 CC FT DOMAIN 5623 5643
 CC FT TRANSMEM 5644 5664
 CC FT DOMAIN 5665 5685
 CC FT TRANSMEM 5686 5706
 CC FT DOMAIN 5707 5727
 CC FT TRANSMEM 5728 5748
 CC FT DOMAIN 5749 5769
 CC FT TRANSMEM 5770 5790
 CC FT DOMAIN 5791 5811
 CC FT TRANSMEM 5812 5832
 CC FT DOMAIN 5833 5853
 CC FT TRANSMEM 5854 5874
 CC FT DOMAIN 5875 5895
 CC FT TRANSMEM 5896 5916
 CC FT DOMAIN 5917 5937
 CC FT TRANSMEM 5938 5958
 CC FT DOMAIN 5959 5979
 CC FT TRANSMEM 5980 6000
 CC FT DOMAIN 6001 6021
 CC FT TRANSMEM 6022 6042
 CC FT DOMAIN 6043 6063
 CC FT TRANSMEM 6064 6084
 CC FT DOMAIN 6085 6105
 CC FT TRANSMEM 6106 6126
 CC FT DOMAIN 6127 6147
 CC FT TRANSMEM 6148 6168
 CC FT DOMAIN 6169 6189
 CC FT TRANSMEM 6190 6210
 CC FT DOMAIN 6211 6231
 CC FT TRANSMEM 6232 6252
 CC FT DOMAIN 6253 6273
 CC FT TRANSMEM 6274 6294
 CC FT DOMAIN 6295 6315
 CC FT TRANSMEM 6316 6336
 CC FT DOMAIN 6337 6356
 CC FT TRANSMEM 6357 6377
 CC FT DOMAIN 6378 6398
 CC FT TRANSMEM 6399 6419
 CC FT DOMAIN 6420 6440
 CC FT TRANSMEM 6441 6461
 CC FT DOMAIN 6462 6482
 CC FT TRANSMEM 6483 6503
 CC FT DOMAIN 6504 6524
 CC FT TRANSMEM 6525 6545
 CC FT DOMAIN 6546 6566
 CC FT TRANSMEM 6567 6587
 CC FT DOMAIN 6588 6608
 CC FT TRANSMEM 6609 6629
 CC FT DOMAIN 6630 6650
 CC FT TRANSMEM 6651 6671
 CC FT DOMAIN 6672 6692
 CC FT TRANSMEM 6693 6713
 CC FT DOMAIN 6714 6734
 CC FT TRANSMEM 6735 6755
 CC FT DOMAIN 6756 6776
 CC FT TRANSMEM 6777 6797
 CC FT DOMAIN 6798 6818
 CC FT TRANSMEM 6819 6839
 CC FT DOMAIN 6840 6860
 CC FT TRANSMEM 6861 6881
 CC FT DOMAIN 6882 6902
 CC FT TRANSMEM 6903 6923
 CC FT DOMAIN 6924 6944
 CC FT TRANSMEM 6945 6965
 CC FT DOMAIN 6966 6986
 CC FT TRANSMEM 6987 7007
 CC FT DOMAIN 7008 7028
 CC FT TRANSMEM 7029 7049
 CC FT DOMAIN 7050 7070
 CC FT TRANSMEM 7071 7091
 CC FT DOMAIN 7092 7112
 CC FT TRANSMEM 7113 7133
 CC FT DOMAIN 7134 7154
 CC FT TRANSMEM 7155 7175
 CC FT DOMAIN 7176 7196
 CC FT TRANSMEM 7197 7217
 CC FT DOMAIN 7218 7238
 CC FT TRANSMEM 7239 7259
 CC FT DOMAIN 7260 7280
 CC FT TRANSMEM 7281 7301
 CC FT DOMAIN 7302 7322
 CC FT TRANSMEM 7323 7343
 CC FT DOMAIN 7344 7364
 CC FT TRANSMEM 7365 7385
 CC FT DOMAIN 7386 7406
 CC FT TRANSMEM 7407 7427
 CC FT DOMAIN 7428 7448
 CC FT TRANSMEM 7449 7469
 CC FT DOMAIN 7470 7490
 CC FT TRANSMEM 7491 7511
 CC FT DOMAIN 7512 7532
 CC FT TRANSMEM 7533 7553
 CC FT DOMAIN 7554 7574
 CC FT TRANSMEM 7575 7595
 CC FT DOMAIN 7596 7616
 CC FT TRANSMEM 7617 7637
 CC FT DOMAIN 7638 7658
 CC FT TRANSMEM 7659 7679
 CC FT DOMAIN 7680 7700
 CC FT TRANSMEM 7701 7721
 CC FT DOMAIN 7722 7742
 CC FT TRANSMEM 7743 7763
 CC FT DOMAIN 7764 7784
 CC FT TRANSMEM 7785 7805
 CC FT DOMAIN 7806 7826
 CC FT TRANSMEM 7827 7847
 CC FT DOMAIN 7848 7868
 CC FT TRANSMEM 7869 7889
 CC FT DOMAIN 7890 7910
 CC FT TRANSMEM 7911 7931
 CC FT DOMAIN 7932 7952
 CC FT TRANSMEM 7953 7973
 CC FT DOMAIN 7974 7994
 CC FT TRANSMEM 7995 8015
 CC FT DOMAIN 8016 8036
 CC FT TRANSMEM 8037 8057
 CC FT DOMAIN 8058 8078
 CC FT TRANSMEM 8079 8099
 CC FT DOMAIN 8100 8120
 CC FT TRANSMEM 8121 8141
 CC FT DOMAIN 8142 8162
 CC FT TRANSMEM 8163 8183
 CC FT DOMAIN 8184 8204
 CC FT TRANSMEM 8205 8225
 CC FT DOMAIN 8226 8246
 CC FT TRANSMEM 8247 8267
 CC FT DOMAIN 8268 8288
 CC FT TRANSMEM 8289 8309
 CC FT DOMAIN 8310 8330
 CC FT TRANSMEM 8331 8351
 CC FT DOMAIN 8352 8372
 CC FT TRANSMEM 8373 8393
 CC FT DOMAIN 8394 8414
 CC FT TRANSMEM 8415 8435
 CC FT DOMAIN 8436 8456
 CC FT TRANSMEM 8457 8477
 CC FT DOMAIN 8478 8498
 CC FT TRANSMEM 8499 8519
 CC FT DOMAIN 8520 8540
 CC FT TRANSMEM 8541 8561
 CC FT DOMAIN 8562 8582
 CC FT TRANSMEM 8583 8603
 CC FT DOMAIN 8604 8624
 CC FT TRANSMEM 8625 8645
 CC FT DOMAIN 8646 8666
 CC FT TRANSMEM 8667 8687
 CC FT DOMAIN 8688 8708
 CC FT TRANSMEM 8709 8729
 CC FT DOMAIN 8730 8750
 CC FT TRANSMEM 8751 8771
 CC FT DOMAIN 8772 8792
 CC FT TRANSMEM 8793 8813
 CC FT DOMAIN 8814 8834
 CC FT TRANSMEM 8835 8855
 CC FT DOMAIN 8856 8876
 CC FT TRANSMEM 8877 8897
 CC FT DOMAIN 8898 8918
 CC FT TRANSMEM 8919 8939
 CC FT DOMAIN 8940 8960
 CC FT TRANSMEM 8961 8981
 CC FT DOMAIN 8982 9002
 CC FT TRANSMEM 9003 9023
 CC FT DOMAIN 9024 9044
 CC FT TRANSMEM 9045 9065
 CC FT DOMAIN 9066 9086
 CC FT TRANSMEM 9087 9107
 CC FT DOMAIN 9108 9128
 CC FT TRANSMEM 9129 9149
 CC FT DOMAIN 9150 9170
 CC FT TRANSMEM 9171 9191
 CC FT DOMAIN 9192 9212
 CC FT TRANSMEM 9213 9233
 CC FT DOMAIN 9234 9254
 CC FT TRANSMEM 9255 9275
 CC FT DOMAIN 9276 9296
 CC FT TRANSMEM 9297 9317
 CC FT DOMAIN 9318 9338
 CC FT TRANSMEM 9339 9359
 CC FT DOMAIN 9360 9380
 CC FT TRANSMEM 9381 9401
 CC FT DOMAIN 9402 9422
 CC FT TRANSMEM 9423 9443
 CC FT DOMAIN 9444 9464
 CC FT TRANSMEM 9465 9485
 CC FT DOMAIN 9486 9506
 CC FT TRANSMEM 9507 9527
 CC FT DOMAIN 9528 9548
 CC FT TRANSMEM 9549 9569
 CC FT DOMAIN 9570 9590
 CC FT TRANSMEM 9591 9611
 CC FT DOMAIN 9612 9632
 CC FT TRANSMEM 9633 9653
 CC FT DOMAIN 9654 9674
 CC FT TRANSMEM 9675 9695
 CC FT DOMAIN 9696 9716
 CC FT TRANSMEM 9717 9737
 CC FT DOMAIN 9738 9758
 CC FT TRANSMEM 9759 9779
 CC FT DOMAIN 9780 9800
 CC FT TRANSMEM 9801 9821
 CC FT DOMAIN 9822 9842
 CC FT TRANSMEM 9843 9863
 CC FT DOMAIN 9864 9884
 CC FT TRANSMEM 9885 9905
 CC FT DOMAIN 9906 9926
 CC FT TRANSMEM 9927 9947
 CC FT DOMAIN 9948 9968
 CC FT TRANSMEM 9969 9989
 CC FT DOMAIN 9990 10000

FT CONFLICT 417 417 S -> T (IN REF. 1).
SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAFB96 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 694;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHOE8TXGRSRGRSGXSGS 19
DB 182 SYTEAGSGSGSGSGSGSGS 200

RESULT 5

FRQ_CRESP STANDARD; PRT; 1015 AA.

AC 000586;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Frequency clock protein.

GN PRO.

OS Creopus spinulosus (Chromocrea spinulosa).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocremycetidae; Hypocreales; Hypocreaceae; Hypocrea.

OX NCBI_TaxID=110619;

RN [1]

PC SEQUENCE FROM N.A.
PC STRAIN=PGSC 4896;

RX MEDLINE=97186515; PubMed=9037100;
RA Lewis M.T., Morgan L.W., Feldman J.F.;

RT "Analysis of frequency (frq) clock gene homologs: evidence for a
RT Mol. Gen. Genet. 253:401-414(1997)."

CC -1- FUNCTION: Circadian clock component involved in the generation of
CC biological rhythms, in particular in rhythm stability, period
CC length, and temperature compensation. Behaves as a negative
CC element in circadian transcriptional loop (by similarity).

CC -1- SIMILARITY: BELONGS TO THE FRQ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U25850; AAA68072.1; -.

DR PIR: T42013; T42013.

KM Biological rhythms; Transcription regulation; Nuclear protein.

FT DOMAIN 240 245 POLY-SER.
FT DOMAIN 356 368 POLY-GLN.

FT DOMAIN 443 451 POLY-PRO.
FT DOMAIN 584 588 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 892 913 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 SHOE8TXGRSRGRSGXSGS 19
DB 583 NHRGKXGHRGSDSGSSGN 601

RESULT 6

SSB TREPA STANDARD; PRT; 176 AA.

AC 083101;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
GN SSB OR TP0062.

OS Treponema pallidum.

OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=160;

RN [1]

PC SEQUENCE FROM N.A.

PC STRAIN=Nichols;

RX MEDLINE=98332770; PubMed=965876;

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,

RA Sodergren E., Hardham U.M., McLeod M.P., Salzberg S., Peterson J.,

RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,

RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,

RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,

RA Venter J.C.;

RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."

RL Science 281:375-388(1998).

CC -1- FUNCTION: This protein is essential for replication of the
CC chromosome. It is also involved in DNA recombination and repair
CC (by similarity).

CC -1- SIMILARITY: Contains 1 SSB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AE001191; AAC65057.1; -.

DR PIR: F71370; F71370.

DR HSSP: P02339; IBYG.

DR TIGR: TP0062; -.

DR InterPro: IPR008994; Nucleic acid DB.

DR InterPro: IPR000424; SSB_protein.

DR Pfam: PF00435; SSB; 1.

DR TIGRFAMs: TIGR00621; ssb; 1.

DR PROSITE: PS50935; SSB; 1.

KM DNA-binding; DNA repair; DNA replication; Complete proteome.

FT DOMAIN 4 106 SSB.

SQ SEQUENCE 176 AA; 19043 MW; 47C49D2F7927D2C CRC64;

Query Match 47.7%; Score 42; DB 1; Length 176;
Best Local Similarity 62.5%; Pred. No. 4.6;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 ESTXGRSRGRSGXSGS 19
DB 128 ESTAGRVRGTSSDSRS 143

RESULT 7

GRP3 ARTSA STANDARD; PRT; 308 AA.

AC P13330;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycine-rich protein GRP33.

OS Artemia salina (Brine shrimp).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;

OC Artemiidae; Artemia.

OX NCBI_TaxID=85549;

RN [1]

PC SEQUENCE FROM N.A.

PC MEDLINE=88007550; PubMed=2443491;

RA Cruz-Alvarez M., Pellicer A.;

RT "Cloning of a full-length complementary DNA for an Artemia salina
RT glycine-rich protein. Structural relationship with RNA binding
RT proteins."

RL J. Biol. Chem. 262:13377-13380(1987).
 CC -1- PTM: THE ARGININES IN THE GLY-RICH DOMAIN MIGHT BE METHYLATED.
 CC -1- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 CC HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 CC STRICTLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC -1- SIMILARITY: Contains 1 KH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03453; AAC03400.1; -.
 DR PIR: A29379; A29379.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR SMART: SM00322; KH_1.
 DR PROSITE: PS50084; KH_type_1, 1.
 DR Nuclear protein; Ribonucleoprotein; Methylation; RNA-binding.
 FT DOMAIN 83 118 KH.
 FT 186 308 GLY-RICH.
 SQ SEQUENCE 308 AA; 33034 MW; E534CE552BB3BC2 CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 308;
 Best Local Similarity 72.7%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 GSRGRSGXSG 18
 DB 211 GGRGRGSGSG 221
 RESULT 8
 SOL2_YEAST STANDARD; PRT; 315 AA.
 ID SOL2_YEAST
 AC F37262;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (PGL).
 GN SOL2 OR YCXX13W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RP [1] SEQUENCE FROM N.A.
 RP RX MEDLINE=96363906; PubMed=8725220;
 RA Shen W.C., Stanford D.R., Hopper A.K.
 RT "Lossip, involved in yeast pre-tRNA splicing, positively regulates
 RT members of the SOL gene family.";
 RL Genetics 143:699-712(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ballesta J.F.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
 RA Sanz E.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION, AND SIMILARITY.
 RP RX MEDLINE=94147996; PubMed=8313894;
 RA Koonin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RL EMBO J. 13:493-503(1994).
 CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
 CC phosphogluconate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
 CC phospho-D-gluconate.
 CC -1- PATHWAY: Pentose phosphate pathway; second step.
 CC -1- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate
 CC isomerase family. 6-phosphogluconolactonase subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U46559; AAB49322.1; -.
 DR EMBL: X59720; CAA42272.1; -.
 DR PIR: S53589; S53589.
 DR Germline, 139004; -.
 DR SGD: S0000718; SOL2.
 DR GO: GO:0008033; P:RNA processing; IGI.
 DR InterPro: IPR006148; Gluc_gal_isom.
 DR InterPro: IPR005900; Phosphogluconac.
 DR Pfam: PF01182; Glucosamine Iso. 1.
 DR TIGRfams: TIGR01198; p91, 1.
 KM Hydrolyase.
 FT CONFLICT 171 171 A -> V (IN REF. 2).
 SQ SEQUENCE 315 AA; 34501 MW; 7F0AD76574AB276 CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 315;
 Best Local Similarity 50.0%; Pred. No. 8.7;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 4 ESTYGRSGRSGXSGS 19
 DB 60 KSTASAEKSGSGSGS 75

RESULT 9
 KRE1_CAEEL STANDARD; PRT; 772 AA.
 ID KRE1_CAEEL
 AC Q23255;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Double-strand break repair protein mre-11.
 GN MRE-11 OR ZC302.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RP [1] SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
 RP GLD-139.
 RP RX MEDLINE=21135651; PubMed=11238374;
 RA Chin G.M., Villeneuve A.M.;
 RT "C. elegans mre-11 is required for meiotic recombination and DNA
 RT repair but is dispensable for the meiotic G(2) DNA damage
 RT checkpoint.";
 RL Genes Dev. 15:522-534(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP RC STRAIN=Bristol N2;
 RA Kelly P.F.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
 CC processing.
 CC -1- FUNCTION: Required for meiotic crossing over and chiasma
 CC formation. Pachytene morphology and homolog pairing are normal.
 CC vital in long term for maintenance of reproductive capacity of
 CC subsequent generations.
 CC -1- COPFACTOR: Manganese (By similarity).
 CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the KRE1/RAD32 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z73978; CAA98292.1; -
 DR PIR: T27512; T27512.
 DR GenBank: 208612; -
 DR WormPep; ZC302.1; CE06573.
 DR InterPro; IPR003701; DNA_repair.
 DR InterPro; IPR004843; M-peptidase.
 DR InterPro; IPR007281; Mre11_DNA_bind.
 DR Pfam; PF00149; Metallophos; 1.
 DR Pfam; PF04152; Mre11_DNA_bind; 1.
 DR TIGRFAMs; TIGR00583; mre11; 1.
 DR DNA_repair; Hydrolase; Nuclease; Endonuclease; Exonuclease;
 KM Nuclear protein; Manganese; Melosis.
 FT MDTAGEN 139 E-SK; IN MRE-11-ME41; DEFECTIVE IN
 FT MEIOTIC CHROMOSOME DEGRADATION.
 SQ SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 772;
 Best Local Similarity 47.4%; Pred. No. 23;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
 QY 1 SHGRTXGRSGRSGSGS 19
 Db 697 SKQPTTRGRGRGARAGAS 715
 RESULT 10
 COAC_YEAST STANDARD; PRT; 2233 AA.
 AC 000955;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin
 DE carboxylase (EC 6.3.4.14)].
 GN FAS3 OR ACC1 OR YNR016C OR N3175.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 2015-2022.
 RX MEDLINE=92262474; PubMed=1350093;
 RA Al-Feel W., Chirala S.S., Wakil S.J.;
 RT "Cloning of the yeast FAS3 gene and primary structure of yeast
 RT acetyl-CoA carboxylase";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Pohl T.M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC FUNCTION: This protein carries three functions: biotin carboxyl
 CC carrier protein, biotin carboxylase, and carboxyltransferase.
 CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
 CC + malonyl-CoA.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carrier protein.
 CC -1- COFACTOR: Biotin.
 CC -1- ENZYME REGULATION: By phosphorylation.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
 CC step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M92156; AAA20073.1; -
 DR EMBL: Z71631; CAA96294.1; -
 DR PIR: S63347; S63347.
 DR HSSD; P24182; IENC.
 DR GenBank: 143361; -
 DR SGD; S0005299; ACC1.
 DR GO; GO:0005789; C:cytoplasmic reticulum membrane; IDA.
 DR GO; GO:0003989; F:acetyl-CoA carboxylase activity; IMP.
 DR GO; GO:0004075; F:biotin carboxylase activity; IMP.
 DR InterPro; IPR001882; Biotin_BS.
 DR InterPro; IPR005482; Biotin_carb_C.
 DR InterPro; IPR000089; Biotin_lipoYL.
 DR InterPro; IPR000022; Carboxyl_trans.
 DR InterPro; IPR005479; Cbase_L_D2.
 DR InterPro; IPR005481; Cbase_L_N.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF00364; Biotin_lipoYL; 1.
 DR Pfam; PF00139; Carboxyl_trans; 1.
 DR Pfam; PF00289; Cbase_L_D2; 1.
 DR Pfam; PF02786; Cbase_L_D2; 1.
 DR PROSITE; PS00189; BIOTIN; 1.
 DR PROSITE; PS00866; CBASE_1; 1.
 DR PROSITE; PS00867; CBASE_2; 1.
 KM Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme;
 KW ATP-binding; Phosphorylation.
 FT NP_BIND 256 261
 FT ACT_SITE 383 383
 FT BINDING 735 735
 FT DOMAIN 1865 1894
 FT CONFLICT 1523 1523
 FT CONFLICT 1755 1755
 FT CONFLICT 1761 1766
 SQ SEQUENCE 2233 AA; 250351 MW; 0A335AAD951F8308 CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 2233;
 Best Local Similarity 50.0%; Pred. No. 72;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 HOESTXGRSGRSGSGS 19
 Db 1200 HQSSNGPAPDRSGSSAS 1217
 RESULT 11
 S55_RAT STANDARD; PRT; 825 AA.
 AC Q6303;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE S55 antigen.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistar; TISSUE=Brain;
 RX MEDLINE=96015159; PubMed=8537300;
 RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
 RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
 RT protein S55 in the nervous system";
 RL J. Biochem. 118:122-128 (1995).
 CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Expressed in neurons.
 CC -----


```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D37934; BAA07153.1; -.
DR PIR: JC4163; JC4163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD55BDF CRC64;

Query Match 46.6%; Score 41; DB 1; Length 825;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GRSRGRSGSGS 19
Db 589 GRSRGRSGSGS 600

RESULT 12
ID ROA3_MOUSE STANDARD; PRT; 379 AA.
AC Q8BG05; Q8BHF8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3).
GN HNRPA3 OR HNRNPA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RA Bayaralhan D.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=FVB/N; TISSUE=Eye;
RA MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Mansura K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Uebelin T.B., Tomshiyki S., Carninci P., Prange C.S.,
RA Rana S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shewchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [1] SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes (By
RN similarity).
RN [2] ALTERNATIVE PRODUCTS:
RN Event=Alternative splicing; Named isoforms=2;
RN Name=1;
RN IsoId=Q8BG05-1; Sequence=Displayed;
RN Name=2;
RN IsoId=Q8BG05-2; Sequence=VSP_007350;
CC -1 SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
CC GYR-RICH PROTEINS.

```

```

CC -1 SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF463524; AAN7692.1; -.
DR EMBL: BC023828; AAH23828.1; -.
DR EMBL: BC023908; AAH23908.1; -.
DR EMBL: BC038364; AAH38364.1; -.
DR InterPro: IPR005054; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS0102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; 2.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
KW Alternative splicing.
FT DOMAIN 35 118 RNA-BINDING (RRM) 1.
FT DOMAIN 126 205 RNA-BINDING (RRM) 2.
FT DOMAIN 211 379 GYR-RICH.
FT VARSPPLIC 1 23 WEKKPPGRPOPDSGRRRRRGE -> M (in isoform
FT 2).
SQ SEQUENCE 379 AA; 39652 MW; D83C400A2B096E9B CRC64;

Query Match 46.0%; Score 40.5; DB 1; Length 379;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 SHQE-STXGRSGRSGSGS 19
Db 203 SKQEWQSGSGRGRSGSGN 222

RESULT 13
ID THO4_MOUSE STANDARD; PRT; 255 AA.
AC Q08583; Q8CBM4; Q9JW7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE THO complex subunit 4 (Tho4) (RNA and export factor binding protein 1)
DE (REF1-1) (Ally of AMD-1 and LEF-1) (Ally/REF).
DE THO4 OR REF1 OR REF1 OR ALY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR LOCATION, AND
RN TISSUE SPECIFICITY.
RN TISSUE=Thymocytes;
RN MEDLINE=97230366; PubMed=9119228;
RN Bruhn L., Munerlyn A., Grosschedl R.;
RT "Ally, a context-dependent coactivator of LEF-1 and AMD-1, is required
RT for TCRAlph enhancer function."
RL Genes Dev. 11:640-653(1997).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, AND INTERACTION WITH NXF1.
RN TISSUE=Embryo;
RN MEDLINE=20246526; PubMed=10786854;
RN Stutz F., Bach A., Doerks T., Braun I.C., Seraphin B., Wilm M.,
RN Bork P., Izaurralde E.;
RT "REF, an evolutionarily conserved family of hnRNP-like proteins,
RT interacts with TAP/Mex57p and participates in mRNA nuclear export."
RL RNA 6:638-650(2000).
RN [3]
RN SEQUENCE OF 190-255 FROM N.A.
RN STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RC

```

RA MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikiado I., Osego N., Saito R., Suzuki H., Yamahara I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gochori T.,
 RA Badaricelli R., Hill D.P., Bult C., Hume D.A., Quakenbush J.,
 RA Schiml L.M., Karapin A., Matsuda H., Batatov S., Betsel K.W.,
 RA Blake U.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Datta B., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
 RA Gasteierland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.T., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedziercki R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nnata K., Okito T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,
 RA Ravey T., Reed J.C., Reed D.J., Reid U., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Sempile C.A., Secou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verrardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan W., Zhu Y., Zimmer A., Carinici P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai U., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RT Nature 420:563-573 (2002).
 [4]
 RA INTERACTION WITH THE EXON JUNCTION COMPLEX.
 RA MEDLINE=22088306; PubMed=12093754;
 RA Lejeune F., Ishigaki Y., Li X., Maquat L.E.;
 RT "The exon junction complex is detected on CBP80-bound but not
 RT eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling.";
 RT EMBO J. 21:3536-3545 (2002).
 [5]
 RA STRUCTURE BY NMR OF 105-182.
 RA MEDLINE=22694206; PubMed=12803490;
 RA Perez-Alvarado G.C., Martinez-Yamout M., Allen M.M., Groschedl R.,
 RA Dyson H.U., Wright P.E.;
 RT "Structure of the nuclear factor Aly: insights into
 RT post-transcriptional regulatory and mRNA nuclear export processes.";
 RT Biochemistry 42:7348-7357 (2003).
 -1- FUNCTION: Acts as a chaperone and promotes the dimerization of
 CC transcription factors containing basic leucine zipper (bZIP)
 CC domains and thereby promotes transcriptional activation. Plays a
 CC role in mRNA processing and export. May function as a scaffold that
 CC mediates interactions between proteins and/or RNA. Integral part
 CC of the THO/TREX complex that is recruited to transcribed genes and
 CC travels with the RNA polymerase during elongation. Is part of the
 CC exon junction complex that remains associated with spliced mRNA
 CC and plays an important role in mRNA export and nonsense-mediated
 CC RNA decay.
 CC SUBUNIT: Homomultimer. Binds LEFT1 and RUNX1. Is part of several
 CC complexes involved in mRNA processing and export. Part of the
 CC heteromultimeric THO/TREX complex containing THO1, THO2, THO3,
 CC THO4 and NXF1/UAP56. Associates with the spliceosome (by
 CC similarity). Binds NXF1, NXF2 and RBM8A directly and is part of
 CC the exon junction complex (EJC) containing NCBP1, NCBP2, NCBP3,
 CC RBM8A, SRRM1, NXF1, NXF2, UPF3B, UPF2, THO4 and/or RBM8A.
 CC SUBCELLULAR LOCATION: Nuclear. Travels to the cytoplasm as part of
 CC the exon junction complex (EJC) bound to mRNA.
 CC ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RefSeq-I;
 CC Name=2; Synonyms=RefSeq-II;
 CC IsoId=008583-1; Sequence=Displayed;
 CC IsoId=008583-2; Sequence=VSP_008597;
 CC TISSUE SPECIFICITY: Highly expressed in heart, brain, spleen,
 CC lung, liver, skeletal muscle, kidney and testis.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; U89876; AAC53117.1; -
 CC EMBL; AJ252140; CAB76383.1; -
 CC EMBL; AK035721; BAC29168.1; -
 CC PDB; 1N08; 12-AUG-03.
 CC TRANSFAC; T03213; -
 CC MGD; MGI:1341044; RefSeq.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00076; rnm; 1.
 CC SMART; SM00360; rnm; 1.
 CC PROSITE; PS50102; RNM; 1.
 CC PROSITE; PS00030; RNM_RNP_1; FALSE_NRG.
 CC Chapterone; Transport; mRNA processing; mRNA splicing;
 CC Nuclear protein; Spliceosome; RNA-binding; Alternative splicing;
 CC 3D-structure.
 CC FT DOMAIN 21 229 GUY/ALA/ARG-RICH.
 CC FT DOMAIN 105 182 RNA-BINDING (RBM).
 CC FT VARSPPLIC 14 105 Missing (in isoform 2).
 CC FT /FTID=VSP_008597.
 CC FT
 CC SQ SEQUENCE 255 AA; 26940 MW; F597235EDD47C17 CEC64;
 CC
 CC Query Match 45.5%; Score 40; DB 1; Length 255;
 CC Best Local Similarity 63.6%; Pred. No. 15;
 CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 8 GRSRGRSGXSG 18
 CC Db 26 GRSRGRSGXSG 36
 CC
 CC RESULT 14
 CC THO4_HUMAN STANDARD; PRT; 257 AA.
 CC ID Q86V81; 043672;
 CC AC Q86V81; 043672;
 CC DT 10-OCT-2003 (Rel. 42, Created)
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE THO complex subunit 4 (Tho4) (Aly of AML-1 and LRF-1)
 CC DE (Transcriptional coactivator Aly/REF) (bZIP enhancing factor BRF).
 CC GN THO4 OR Aly OR BRF.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCB1_TaxId=9606;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=lung;
 CC RX MEDLINE=22388257; PubMed=12477932;
 CC RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 CC RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 CC RA Diatchenko L., Marnusina K., Farmer A.A., Rubin G.M., Hong L.,
 CC RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 CC RA Brownstein M.J., Usdin T.B., Toshitoki S., Carinici P., Prange C.,
 CC RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 CC RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 CC RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC RA Fahy T., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 CC RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 CC RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 CC RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT [2].
RP SEQUENCE OF 25-257 FROM N.A., AND DISEASE.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=99135498; PubMed=5952027;
RA Wichmann I., Garcia-Lozano J.R., Respalda N.,
RA Gonzalez-Becirban M.F., Munoz-Roldan A.;
RT "Antibodies to transcriptional regulation proteins DEX and ALY in
RT a patient with systemic lupus erythematosus.";
RL Hum. Immunol. 60:57-62(1999).
RN [3].
RP SEQUENCE OF 127-131, 182-189, 239-245 AND 249-251, FUNCTION, AND
RP SUBCELLULAR LOCATION.
RX MEDLINE=99417961; PubMed=10488337;
RA Virasius C.-M., Wagner S., Green M.R.;
RT "A human nuclear-localized chaperone that regulates dimerization, DNA
RT binding, and transcriptional activity of bZIP proteins.";
RL Mol. Cell 4:219-228(1999).
RN [4].
RP PARTIAL SEQUENCE, MASS SPECTROMETRY, AND INTERACTION WITH THE
RP SPLICEOSOME.
RX MEDLINE=98400255; PubMed=9731529;
RA Neubauer G., King A., Rappalber J., Calvio C., Watson M., Ajuh P.,
RA Sleeman J., Lamond A., Mann M.;
RT "Mass spectrometry and EST-database searching allows characterization
RT of the multi-protein spliceosome complex.";
RL Nat. Genet. 20:46-50(1998).
RN [5].
RP SEQUENCE OF 108-133, MASS SPECTROMETRY, AND INTERACTION WITH THE
RP SPLICEOSOME.
RX MEDLINE=22166312; PubMed=12176931;
RA Rappalber J., Ryder U., Lamond A.I., Mann M.;
RT "Large-scale proteomic analysis of the human spliceosome.";
RL Genome Res. 12:1231-1245(2002).
RN [6].
RP FUNCTION, AND INTERACTION WITH RBM8A, NXF1 AND THE EXON JUNCTION
RP COMPLEX.
RX MEDLINE=21564074; PubMed=11707413;
RA Katakoka N., Diem M.D., Kim V.N., Yong J., Dreyfuss G.;
RT "Mago, a human homolog of Drosophila mago nashi protein, is a
RT component of the splicing-dependent exon-exon junction complex.";
RL EMBO J. 20:6424-6433(2001).
RN [7].
RP FUNCTION, AND INTERACTION WITH NXF1 AND NXF1.
RX MEDLINE=21530486; PubMed=11675789;
RA Luo M.-L., Zhou Z., Magni K., Christoforides C., Rappalber J.,
RA Mann M., Reed R.;
RT "Pre-mRNA splicing and mRNA export linked by direct interactions
RT between UAP56 and Aly.";
RL Nature 413:644-647(2001).
RN [8].
RP FUNCTION, AND INTERACTION WITH ULS4.
RX MEDLINE=22326840; PubMed=12438613;
RA Chen I.-H.B., Schlabda K.S., Sandri-Goldin R.M.;
RT "ICP27 interacts with the RNA export factor Aly/REF to direct herpes
RT simplex virus type 1 intronless mRNAs to the TAP export pathway.";
RL J. Virol. 76:12877-12889(2002).
RN [9].
RP FUNCTION, AND INTERACTION WITH THE TREX COMPLEX.
RX MEDLINE=22010388; PubMed=11979277; Pfannstiel J., Opijz M.,
RA Straesser K., Masuda S., Mason P., Pfannstiel J., Opijz M.,
RA Rodriguez-Navarro S., Rondon A.G., Aguilera A., Struhl K., Reed R.,
RA Hurt E.;
RT "TREX is a conserved complex coupling transcription with messenger
RT RNA export.";
RL Nature 417:304-308(2002).
RN [10].
RP FUNCTION: Acts as chaperone and promotes the dimerization of
RP transcription factors containing basic leucine zipper (bZIP)
RP domains and thereby promotes transcriptional activation. Plays a
RP role in mRNA processing and export. May function as scaffold that
RP mediates interactions between proteins and/or RNA. Integral part

CC of the THO/TREX complex that is recruited to transcribed genes and
CC travels with the RNA polymerase during elongation. Is part of the
CC exon junction complex that remains associated with spliced mRNA
CC and plays an important role in mRNA export and nonsense-mediated
CC RNA decay. Directs mRNA derived from Herpes simplex virus intron-
CC less genes to the NXF1-mediated export pathway.
CC
CC - SUBUNIT: Homodimer. Binds REF1 and RUVX1 (By similarity). Is
CC part of several complexes involved in mRNA processing and export.
CC Part of the heteromultimeric THO/TREX complex containing THOC1,
CC THOC2, THOC3, THOC4 and NXF1/UAP56. Associates with the
CC spliceosome. Binds NXF1 and RBM8A directly and is part of
CC the exon junction complex (EJC) containing NCBP1, NCBP2, ENPS1,
CC RBM8A, SRRM1, NXF1, NXF1, UPF3B, UPF2 and THOC4 (By similarity).
CC Binds ULS4 from Herpes simplex type 1 (strain KOS).
CC
CC - SUBCELLULAR LOCATION: Nuclear. Travels to the cytoplasm as part of
CC the exon junction complex (EJC) bound to mRNA.
CC
CC - DISEASE: Antibodies against THOC4 are found in sera of patients
CC with systemic lupus erythematosus (SLE).
CC
CC - SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL, BC052302; AAS52302.1; -
CC EMBL, AF047002; AAD09608.1; -
CC Genew; HGNC:19071; THOC4.
CC MIM; 604171; -
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF000076; rtm; 1.
CC SMART; SM00360; RNM; 1.
CC PROSITE; PS50102; RNM_1; FALSE NEG.
CC DR PROSITE; PS00030; RRM; 1.
CC DR Antigen; Chaperone; Transport; mRNA transport; mRNA processing;
CC KM RNA splicing; Nuclear protein; Spliceosome; RNA-binding.
CC FT DOMAIN 21 231
CC FT 106 183 RNM-BINDING (RNM).
CC FT CONFLICT 26 26 G -> R (IN REF. 2).
CC FT CONFLICT 35 36 G -> R (IN REF. 2).
CC FT CONFLICT 39 39 D -> N (IN REF. 2).
CC FT CONFLICT 150 150 V -> F (IN REF. 2).
CC FT CONFLICT 169 169
CC SO SEQUENCE 257 AA; 26888 MW; E2B5021DA579919A CRC64;
CC
CC Query Match 45.5%; Score 40; DB 1; Length 257;
CC Best Local Similarity 63.6%; Pred. No. 15;
CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 8 GSRGRSGXSG 18
CC DB 26 GRGRGRASQG 36
CC
CC RESULT 15
CC PEXD_PICPA STANDARD; PRT; 380 AA.
CC AC 092266;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Peroxisomal membrane protein PEX13 (Peroxin-13).
CC GN PEX13.
CC OS Pichia pastoris (Yeast).
CC CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC CC Saccharomycetales; Saccharomycetaceae; Pichia.
CC OK NCB1_taxid=4922;
CC RN (1)
CC RP SEQUENCE FROM N.A., AND MUTAGENESIS.
CC MEDLINE=97011155; PubMed=8858165;

RA Gould S.J., Kalish J.E., Morrell J.C., Bjorkman J., Urquhart A.J.,
RT Crane D.I.;
RT "Pex13p is an SH3 protein of the peroxisome membrane and a docking
RT factor for the predominantly cytoplasmic Pts1 receptor.";
RL J. Cell Biol. 135:85-95(1996).
CC -1- FUNCTION: Component of the peroxisomal translocation machinery
CC with PEX14 and PEX17. Interacts with the Pts1 receptor
CC (PAS10/PEX5). Involved in the import of Pts1 and Pts2 proteins.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=-2 isoforms, PEX13L (shown here) and PEX13S, are produced
CC by alternative initiation;
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U70067; AAB09087.1; -.
DR HSSP; Q64010; ICKA.
DR InterPro; IPR007223; Peroxin-13_N.
DR InterPro; IPR001452; SH3.
DR Pfam; PF04088; Peroxin-13_N; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRODOM; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Peroxisome; Transport; Protein transport; Transmembrane; SH3 domain;
KW Alternative initiation.
KW CHAIN 1 380
FT FT CHAIN 33 380
FT FT INIT MET 33 33
FT FT DOMAIN 1 230
FT FT TRANSMEM 231 251
FT FT DOMAIN 252 380
FT FT DOMAIN 277 344
FT FT MUTAGEN 286 286
FT FT MUTAGEN 287 287
FT FT MUTAGEN 288 288
FT FT MUTAGEN 291 291
FT FT MUTAGEN 296 296
SQ SEQUENCE 380 AA; 40695 MW; EEBAC39F93BA832 CRC64;
Query Match 45.5%; Score 40; DB 1; Length 380;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
CY 1 SHQSTXGRSRGRSGXSG 18
DB 315 SNOESTWKKCRSRDGYG 332

Search completed: September 28, 2004, 06:05:52
Job time : 8.22396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-9

Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 segs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriaph:*
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	797	4	Q16824
2	80	90.9	990	4	Q15206
3	80	90.9	1218	4	Q05331
4	78	88.6	798	4	Q9H4U3
5	78	88.6	1084	4	Q01212
6	69	78.4	591	4	Q01720
7	64	72.7	687	4	Q9H4U2
8	63	71.6	322	4	Q03838
9	63	71.6	465	4	Q03838
10	49	55.7	820	12	Q69375
11	49	55.7	822	3	Q9P312
12	47	53.4	436	16	Q82C67
13	46	52.3	476	12	Q80890
14	46	52.3	810	10	Q94E50
15	46	52.3	1142	13	Q7SZV2
16	44	50.0	285	5	Q9W3V3

17	44	50.0	691	4	Q96RC0	Q96RG0 homo sapien
18	44	50.0	722	11	Q08679	Q08679 ratu
19	44	50.0	745	4	Q15524	Q15524 homo sapien
20	44	50.0	755	4	Q96HB3	Q96HB3 homo sapien
21	44	50.0	787	5	Q9GSD4	Q9GSD4 plasmodium
22	44	50.0	888	11	Q8BR95	Q8BR95 mus muscula
23	44	50.0	936	16	Q82H71	Q82H71 streptomyc
24	44	50.0	1284	5	Q9V9Y3	Q9V9Y3 caenorhabdi
25	44	50.0	1686	5	Q21948	Q21948 drosophila
26	44	50.0	1765	5	Q8SX63	Q8SX63 drosophila
27	44	50.0	2232	5	Q9VC96	Q9VC96 drosophila
28	43	48.9	306	5	Q45396	Q45396 caenorhabdi
29	43	48.9	611	10	Q9S767	Q9S767 oryza sativ
30	43	48.9	652	5	Q77168	Q77168 apis mellif
31	43	48.9	756	5	Q9U116	Q9U116 drosophila
32	43	48.9	802	10	Q8S2A7	Q8S2A7 oryza sativ
33	43	48.9	813	10	Q8VZ18	Q8VZ18 arabidopsis
34	43	48.9	843	10	Q9LNX5	Q9LNX5 arabidopsis
35	43	48.9	849	10	Q9C7M2	Q9C7M2 arabidopsis
36	43	48.9	849	10	Q949M6	Q949M6 arabidopsis
37	42.5	48.3	1080	12	Q91AM0	Q91AM0 beet chloro
38	42	47.7	188	10	Q8H8P6	Q8H8P6 oryza sativ
39	42	47.7	199	10	Q9FW52	Q9FW52 oryza sativ
40	42	47.7	240	16	Q87YL4	Q87YL4 bos taurus
41	42	47.7	297	6	Q8S024	Q8S024 bos taurus
42	42	47.7	322	16	Q8G4E5	Q8G4E5 bifidobacte
43	42	47.7	336	10	Q943T8	Q943T8 oryza sativ
44	42	47.7	481	16	Q7UJ99	Q7UJ99 rhodospirell
45	42	47.7	507	2	Q9RMT5	Q9RMT5 streptomyc

ALIGNMENTS

Q16824	PRELIMINARY;	PRT;	797 AA.
AC Q16824;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE Profilaggrin (Fragment).			
GN FLG.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=91064347; PubMed=2248957;			
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Seimert P.M.,			
RT "Organization, structure, and polymorphisms of the human profilaggrin			
RT gene [published erratum appears in Biochemistry 1991 Jun			
RT 11:30(23):5814]";			
RL Biochemistry 29:9432-9440(1990).			
DR EMBL; M60502; AAA63248.1;			
DR GO; GO:0005198; F:structural molecule activity; IEA.			
DR InterPro; IPR003303; Filaggrin.			
DR Pfam; PF03516; Filaggrin; 4.			
DR PRINTS; PRO0487; FILAGRIN.			
FT NON TER			
FT SEQUENCE			
QY	Query Match	95.5%;	Score 84; DB 4; Length 797;
DB	Best Local Similarity	89.5%;	Pred. No. 1.3e-05;
	Matches 17; Conservative	0;	Mismatches 2; Indels 0;
			Gaps 0;
	1 SHQESTXGRSGRSGSGS 19		
	427 SHQESTXGRSGRSGSGS 445		
RESULT 2			

```

O15206
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene [published extractum appears in Biochemistry 1991 Jun
11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AA63244.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 90.9%; Score 80; DB 4; Length 990;
Best Local Similarity 84.2%; Pred. No. 7.8e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 551 SHOESTGRSGRSGSGS 569

RESULT 3
O05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC - FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC - PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDER COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC CC -1- PROTEOLYTICALLY CLEAVED.
CC CC POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1; -
DR FIR; A48118; A48118.
DR HSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CapS 5100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CAP; 1.
KM Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
KM Polymorphism.
FT CA BIND 19 32 SITE I (BY SIMILARITY).
FT CA BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 90.9%; Score 80; DB 4; Length 1218;
Best Local Similarity 84.2%; Pred. No. 9.8e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 773 SHOESTGRSGRSGSGS 791

RESULT 4
O9H4U3 PRELIMINARY; PRT; 798 AA.
AC O9H4U3;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL356504; CAC13171.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 88.6%; Score 78; DB 4; Length 798;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 428 SHOESTGRSGRSGSGS 446

RESULT 5
O01212 PRELIMINARY; PRT; 1084 AA.
ID Q01212

```

```

AC Q01212; Q03840;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-UN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published structure appears in Biochemistry 1991 Jun
RT 11;30(23):5814].".
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin.
DR PRINTS; PR00487; Filaggrin.
DR NON TER
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 88.6%; Score 78; DB 4; Length 1084;
Best Local Similarity 84.2%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19
DB 64 SHOESTRGRSGRSGSGS 82
|||||
Q01720 PRELIMINARY; PRT; 591 AA.
ID Q01720;
AC Q01720;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Nimsukskiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
RL J. Biol. Chem. 267:23772-23781(1992).
DR EMBL; M60503; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR POLYMORPHISM.
-1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
TERMINAL DIFFERENTIATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
-1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
PROTEOLYTICALLY CLEAVED.
-1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
REPEATS.
EMBL; L01089; AAA60177.1; -.
EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR POLYMORPHISM.
KW PROPEP 1 293
FT CHAIN 294 467 POTENTIAL.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FILAGGRIN.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.0032;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19
DB 449 SHOESTRGRSGRSGSGS 467
|||||
Q094U2 PRELIMINARY; PRT; 687 AA.
ID Q094U2;
AC Q094U2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.;
RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.0037;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 SHOESTXGRSGRSGSGS 19
 DB 449 SHOESTXGRSGRSGSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003839;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE FLAGGRIN (PROFLAGGRIN) (Fragment).
 GN Flg.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 29:9432-9440(1990).
 RN (2)
 RP REVISIONS.
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FLAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
 CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
 CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- POLYMORPHISM: A NUMBER OF PROFLAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
 CC REPEATS.
 CC -1- MISCELLANEOUS: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
 CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
 CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSED AS KERATOHYALIN
 CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 DR EMBL: M60499; AAA63246.1; -;
 DR GO: G0:0005198; P:structural molecule activity; IEA.
 DR InterPro: IPR003303; Flaggrin.
 DR Pfam: PF03516; Flaggrin; 3.
 DR PRINTS: PRO0487; FLAGGRIN.
 FT NON_TER 1
 FT 465
 SQ SEQUENCE 465 AA; 50280 MW; C883744C5B134097 CRC64;

Query Match 72.7%; Score 64; DB 4; Length 465;
 Best Local Similarity 73.7%; Pred. No. 0.017;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 SHOESTXGRSGRSGSGS 19
 DB 227 SHOESTXGRSGRSGSGS 245

RESULT 9
 ID 075370 PRELIMINARY; PRT; 322 AA.

AC 075370;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Epidermal filaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OK NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99101527; PubMed=9886436;
 RA Girbal-Neuhauser E., Durieux J.J., Arnaud M., Dalbon P., Sebbaq M.,
 RA Vincent C., Simon M., Sensu T., Masson-Bessiere C.,
 RA Jolivet-Reyraud C., Jolivet M., Serre G.;
 RT "The epitopes targeted by the rheumatoid arthritis-associated
 RT antifilaggrin autoantibodies are posttranslationally generated on
 RT various sites of (pro)filaggrin by deamination of arginine residues."
 RL J. Immunol. 162:585-594(1999).
 DR EMBL: AF043380; AAC23559.1; -;
 DR GO: G0:0005198; P:structural molecule activity; IEA.
 DR InterPro: IPR003303; Flaggrin.
 DR Pfam: PF03516; Flaggrin; 2.
 DR PRINTS: PRO0487; FLAGGRIN.
 FT NON_TER 1
 FT 322
 SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

Query Match 71.6%; Score 63; DB 4; Length 322;
 Best Local Similarity 72.2%; Pred. No. 0.017;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 SHOESTXGRSGRSGSGS 18
 DB 305 SHOESTXGRSGRSGSGS 322

RESULT 10
 ID 069375 PRELIMINARY; PRT; 820 AA.
 AC 069375;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative U125 protein.
 OS Mouse cytomegalovirus 1.
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OK NCBI_TaxID=10366;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K181;
 RX MEDLINE=94233727; PubMed=7513920;
 RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;
 RT "Identification and characterization of a murine cytomegalovirus gene
 RT with homology to the U125 open reading frame of human
 RT cytomegalovirus.";
 RT Virology 200:643-650(1994).
 DR EMBL: U02500; AAA19449.1; -;
 DR InterPro: IPR006731; Herpes_pp85.
 DR Pfam: PF04637; Herpes_pp85; 1.
 SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;

Query Match 55.7%; Score 49; DB 12; Length 820;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 QESTXGRSGRSGS 14
 DB 186 QSTAGRSGRSGS 197

RESULT 11
 ID 099312 PRELIMINARY; PRT; 822 AA.

AC 099312;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.
 GN B12P1.10.


```

OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OK NCBI_TaxID=5141;
RN (1)
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN (12)
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL390091; CAB98213.1; -
DR PIR: T51049; T51049.
DR GO: 0003676; P: nucleic acid binding; IEA.
DR InterPro: IPR000504; RNA_res_mot.
DR InterPro: IPR005120; Smg-4_UPF3.
DR Pfam: PF00076; rrm; 1.
DR Pfam: PF03467; Smg4_UPF3; 1.
DR SMART: SMO0360; RRM; 1.
DR PROSITE: PS00102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 55.7%; Score 49; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGRSGSGSGS 19
DB 414 RESAGSRTGRGRGRT 430

RESULT 12
082C67 PRELIMINARY; PRT; 436 AA.
ID 082C67;
AC 082C67;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative integral membrane protein.
GN SAV5487
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RL MEDLINE=21477403; Pubmed=11572948;
RX Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN (12)
RP SEQUENCE FROM N.A.
RA STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; Pubmed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL: AP005043; BAC73199.1; -
DR InterPro: IPR002194; Capemohn_TCP-1.
DR PROSITE: PS00995; TCP1_3; 1.
RW Complete proteome.
SQ SEQUENCE 436 AA; 45359 MW; 35BD46B80C7282D0 CRC64;

```

```

Query Match 53.4%; Score 47; DB 16; Length 436;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GRSRGRSGSGS 19
DB 253 GRSRGRSGSGS 264

RESULT 13
080890 PRELIMINARY; PRT; 476 AA.
ID 080890;
AC 080890;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE EBNA1.
DE Herpesvirus papio.
OS Vitruces, dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10394;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=594-S;
RX MEDLINE=96400106; Pubmed=8806482;
RA Yates J.L., Camilo S.M., Ali S., Ying A.;
RT "Comparison of the EBNA1 proteins of Epstein-Barr virus and
RT herpesvirus papio in sequence and function.";
RL Virology 222:1-13(1996).
DR EMBL: U23857; AAA66373.1; -
DR HSP: P03211; IYH1.
DR GO: 0042025; C: host cell nucleus; IEA.
DR GO: 0003677; P: DNA binding; IEA.
DR InterPro: IPR004186; EBNA1.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF02905; EBNA1; 1.
SQ SEQUENCE 476 AA; 46082 MW; 3F598F27919F7DA9 CRC64;

Query Match 52.3%; Score 46; DB 12; Length 476;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGSGS 19
DB 162 GRSRGRSGSGS 173

RESULT 14
094ES0 PRELIMINARY; PRT; 810 AA.
ID 094ES0;
AC 094ES0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AAA-metalloprotease FtsH.
GN FtsH.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN (1)
RP SEQUENCE FROM N.A.
RA Kolaczowska A., Kleczawa J., Janska H.;
RT "Pea mitochondrial FtsH can substitute for the yeast m-AAA protease in
RT maintaining the respiratory function.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF397903; AAK77908.1; -
DR GO: 0016020; C: membrane; IEA.
DR GO: 0005524; F: ATP binding; IEA.
DR GO: 0004222; F: metalloendopeptidase activity; IEA.
DR GO: 0000166; F: nucleic acid binding; IEA.
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.

```

DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_centra.
 DR InterPro; IPR003960; AAA_sub.
 DR InterPro; IPR005936; Peptidase_Fish.
 DR InterPro; IPR006642; Peptidase_M1.
 DR Pfam; PF00004; AAA; 1.
 DR Pfam; PF01434; Peptidase_M1; 1.
 DR SMART; SMO0382; AAA; 1.
 DR TIGRFAMs; TIGR01241; Fish_fam; 1.
 DR PROSITE; PS00674; AAA; 1.
 KW ATP-binding; Metalloprotease; Protease.
 SQ SEQUENCE 810 AA; 89317 MW; 22A0EC7ADEB6C18 CRC64;

Query Match 52.3%; Score 46; DB 10; Length 810;
 Best Local Similarity 75.0%; Pred. No. 33;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 8 GRSRGRSGSGS 19
 Db 429 GRKRGCGGSGS 440

RESULT 15

ID 07SZV2 PRELIMINARY; PRT; 1142 AA.
 AC 07SZV2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Period 1.
 GN PERL.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N. A.
 RA Constance C.M., Sulton B., Taylor M., Green C.B.;
 RT "Xenopus laevis period 1.", EMBL/GenBank/DBJ databases.
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY340103; AAC18143.1; -
 SQ SEQUENCE 1142 AA; 123519 MW; 158ECB698DB3E0B6 CRC64;

Query Match 52.3%; Score 46; DB 13; Length 1142;
 Best Local Similarity 58.8%; Pred. No. 47;
 Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Oy 1 SHQSTYGRSRGRSGS 17
 Db 805 SKGRNSGRRRRGKSGS 821

Search completed: September 28, 2004, 06:12:48
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-9
Perfect score: 88
Sequence: 1 SHQESTXGRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.290a04:*
1: geneseqp19808:*
2: geneseqp19905:*
3: geneseqp20008:*
4: geneseqp20015:*
5: geneseqp20025:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	96.6	19	2	AAW61515 Peptide c
2	84	95.5	19	2	AAW61517 Peptide c
3	84	95.5	19	2	AAW61509 Peptide c
4	84	95.5	19	2	AAW61513 Peptide c
5	84	95.5	19	2	AAW61505 Peptide c
6	84	95.5	19	2	AAW61514 Peptide c
7	84	95.5	19	2	AAW61516 Peptide c
8	80	90.9	14	5	ABW78605 Novel hum
9	78	88.6	19	2	AAW61506 Peptide c
10	78	88.6	19	2	AAW61508 Peptide c
11	78	88.6	19	2	AAW61507 Peptide c
12	78	88.6	19	2	AAW61512 Peptide c
13	78	88.6	19	2	AAW61511 Peptide c
14	78	88.6	19	2	AAW61510 Peptide c
15	75	85.2	21	2	AAW61520 Peptide X
16	75	85.2	22	4	AAW61520 Peptide X
17	58	65.9	330	2	AAW61525 Human fil
18	58	65.9	330	2	AAW61525 Human fil
19	58	65.9	330	2	AAW61525 Human fil
20	54	61.4	330	2	AAW61525 Human fil
21	54	61.4	330	2	AAW61525 Human fil
22	50	56.8	477	6	ABW78614 Novel hum
23	45	51.1	1711	4	AAW79819 Human pro
24	44	50.0	1951	4	AAW78835 Human pro
25	44	50.0	285	4	ABW69359 Drosophila

26	44	50.0	532	4	ABW67173	AbW67173 Drosophila
27	44	50.0	691	5	AAW19050	AAW19050 Human PAR
28	44	50.0	691	6	ABW43965	ABW43965 Human ser
29	44	50.0	722	6	ABW43965	ABW43965 Rat serin
30	44	50.0	724	5	AAW19051	AAW19051 Human PAR
31	44	50.0	724	6	ABW08527	ABW08527 Human ser
32	44	50.0	724	6	ABW43964	ABW43964 Human ser
33	44	50.0	731	7	ADW39074	ADW39074 Human KPP
34	44	50.0	745	6	AAW33550	AAW33550 Human mic
35	44	50.0	745	6	ABW43970	ABW43970 Serine/th
36	44	50.0	745	7	ADW34768	ADW34768 Human ser
37	44	50.0	768	4	AAW78349	AAW78349 Human pro
38	44	50.0	774	6	ABW43969	ABW43969 Serine/th
39	44	50.0	777	6	ABW43968	ABW43968 Human ser
40	44	50.0	780	6	ABW43967	ABW43967 Human ser
41	44	50.0	796	6	ABW11830	ABW11830 Human MDD
42	44	50.0	804	4	AAW79333	AAW79333 Human pro
43	44	50.0	1129	4	ABW58274	ABW58274 Drosophila
44	43.5	49.4	399	6	ABW70654	ABW70654 Recombina
45	43	48.9	569	4	AAW79339	AAW79339 Human pro

ALIGNMENTS

RESULT 1
ID AAW61515
AAW61515 standard; peptide; 19 AA.
AC AAW61515;
XX 26-OCT-1998 (first entry)
DE Peptide cFA, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
OS (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
PN Van Venrooij JMW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI, 1998-398613/34.
DR Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
PS Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX Sequence 19 AA;
XX SQ

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 3.9e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 2
 AAW61517 standard; peptide; 19 AA.

AC AAW61517;

DE 26-OCT-1998 (first entry)

XX Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.7e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 3
 AAW61509 standard; peptide; 19 AA.
 AC AAW61509;

XX 26-OCT-1998 (first entry)
 DE Peptide cfc5, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Modified-site 16 /note="Citruiline"

PN W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 5.7e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 4
 AAW61513

ID AAW61513 standard; peptide; 19 AA.

AC AAW61513;

DE 26-OCT-1998 (first entry)

XX Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Modified-site 7
FT Modified-site /note= "Citruiline"
FT Modified-site 16
FT Modified-site /note= "Citruiline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-M61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-07;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTXGRSGRSGXSGS 19
XX 1 SHOESTXGRSGRSGXSGS 19
XX
XX Db
XX
XX RESULT 5
XX AAM61505
XX ID AAM61505 standard; peptide; 19 AA.
XX
XX AC AAM61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cft, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7
XX /note= "Citruiline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX

```

```

PF 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-M61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 5.7e-07;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTXGRSGRSGXSGS 19
XX 1 SHOESTXGRSGRSGXSGS 19
XX
XX Db
XX
XX RESULT 6
XX AAM61514
XX ID AAM61514 standard; peptide; 19 AA.
XX
XX AC AAM61514;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cf, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7
XX /note= "Citruiline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX

```

AAW61506;
AC
XX

DT 26-OCT-1998 (first entry)
 XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 9 /note="Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 CC
 XX Disclosure; Page 6; 19pp; English.
 PS
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTGRSGRSGSGS 19
 DB 1 SHOESTGRSGRSGRSGS 19
 RESULT 10
 AAW61508
 ID AAW61508 standard; peptide; 19 AA.
 XX
 AC AAW61508;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc4, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 13 /note="Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 CC
 XX Disclosure; Page 6; 19pp; English.
 PS
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTGRSGRSGSGS 19
 DB 1 SHOESTGRSGRSGRSGS 19
 RESULT 11
 AAW61507
 ID AAW61507 standard; peptide; 19 AA.
 XX
 AC AAW61507;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 11 /note="Citruilline"
 FT
 PN WO9822503-A2.
 PD 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
DR
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 84.2%; Pred. No. 5.6e-06;
XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SHOESTXGRSRGRSGSGS 19
DB 1 SHOESTXGRSRGRSGSGS 19
XX
XX RESULT 12
XX AAW61512
XX ID AAW61512 standard; peptide; 19 AA.
XX
XX AAW61512;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7 /note= "Citruilline"
XX Modified-site 13 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 28-MAY-1998.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 5.6e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTXGRSRGRSGSGS 19
DB 1 SHOESTXGRSRGRSGSGS 19
XX
XX RESULT 13
XX AAW61511
XX ID AAW61511 standard; peptide; 19 AA.
XX
XX AAW61511;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7 /note= "Citruilline"
XX Modified-site 11 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC

```


CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGRSGSGS 19
 |||||
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 14
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

AC AAW61510;
 XX 26-OCT-1998 (first entry)

DE Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"

FT Modified-site 9 /note= "Citruilline"

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGRSGSGS 19
 |||||
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 XX 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 1.9e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HQESTXGRSRGRSGSGS 19
 |||||
 DB 4 HQESTXGRSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:20
 Job time : 56.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-9
Perfect score: 88
Sequence: 1 SHQSTXGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 segs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-5
4	84	95.5	19	9	US-09-308-150-9
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	84	95.5	19	9	US-09-308-150-2
9	78	88.6	19	9	US-09-308-150-3
10	78	88.6	19	9	US-09-308-150-4
11	78	88.6	19	9	US-09-308-150-6
12	78	88.6	19	9	US-09-308-150-7
13	78	88.6	19	9	US-09-308-150-8
14	75	85.2	21	9	US-09-308-150-10
15	75	85.2	22	9	US-09-747-029A-22

16	56.8	477	15	US-10-161-927-62	Sequence 62, Appl
17	53.4	123	16	US-10-767-701-32436	Sequence 32436, A
18	53.4	436	14	US-10-156-761-13022	Sequence 13022, A
19	53.4	838	16	US-10-437-963-146503	Sequence 146503, A
20	53.4	854	12	US-10-425-114-57838	Sequence 57838, A
21	52.3	78	16	US-10-437-963-203073	Sequence 203073, A
22	51.1	653	9	US-09-746-801A-13	Sequence 13, Appl
23	50.0	262	12	US-10-425-114-72166	Sequence 72166, A
24	50.0	262	12	US-10-425-114-72167	Sequence 72167, A
25	50.0	563	16	US-10-437-963-198539	Sequence 198539, A
26	50.0	691	44	US-09-919-585-6	Sequence 6, Appl
27	50.0	722	15	US-10-274-184-4	Sequence 4, Appl
28	50.0	722	16	US-10-760-407-4	Sequence 4, Appl
29	50.0	724	9	US-09-919-585-9	Sequence 9, Appl
30	50.0	724	15	US-10-274-194-2	Sequence 2, Appl
31	50.0	724	16	US-10-760-407-2	Sequence 2, Appl
32	50.0	745	12	US-10-260-708-79	Sequence 79, Appl
33	50.0	745	14	US-10-195-101-16	Sequence 16, Appl
34	50.0	745	14	US-10-161-565-24	Sequence 24, Appl
35	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
36	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
37	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
38	48.9	91	16	US-10-437-963-132029	Sequence 132029, A
39	48.9	130	16	US-10-437-963-174101	Sequence 174101, A
40	48.9	134	12	US-10-425-114-43265	Sequence 43265, A
41	48.9	134	12	US-10-425-114-63280	Sequence 63280, A
42	48.9	182	12	US-10-424-599-214180	Sequence 214180, A
43	48.9	500	16	US-10-437-963-190244	Sequence 190244, A
44	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
45	48.9	611	16	US-10-437-963-149539	Sequence 149539, A

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patient No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoel, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.7e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHQSTXGRSRGRSGXSGS 19

Db 1 SHQSTXGRSGRSGSGS 19

RESULT 2

US-09-308-150-1
Sequence 1, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 3

US-09-308-150-5
Sequence 5, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 4

US-09-308-150-9

Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 5

US-09-308-150-11

Sequence 11, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 19

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-11
```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8.1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTGRSRGRSGSGS 19
```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-13
```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8.1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTGRSRGRSGSGS 19
```

```

RESULT 7
US-09-308-150-14
```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-14
```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8.1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTGRSRGRSGSGS 19
```

```

RESULT 8
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-2
```

Query March 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 6,4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 9
US-09-308-150-3

Sequence 3, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-3

Query March 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 6,4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 10
US-09-308-150-4

Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-4

Query March 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6,4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 11
US-09-308-150-6

Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-6

Query March 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6,4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 12
US-09-308-150-7

Sequence 7, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTXGRSGRSGSGS 19

```

```

RESULT 13
US-09-308-150-8
; Sequence 8, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTXGRSGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.0002;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2  HOESTXGRSGRSGSGS 19
Db      4  HOESTXGRSGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NPU500 INNS:031
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-9

Sequence: 1 SHOEYXGRSGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	51.1	653	US-09-513-057C-13	Sequence 13, Appl
2	44	50.0	722	US-09-984-890-4	Sequence 4, Appl
3	44	50.0	724	US-09-984-890-2	Sequence 2, Appl
4	44	50.0	745	US-09-523-849-36	Sequence 36, Appl
5	42	47.7	169	US-09-342-084-6	Sequence 6, Appl
6	42	47.7	448	US-09-252-991A-24066	Sequence 24066, A
7	42	47.7	722	US-08-817-832B-32	Sequence 32, Appl
8	42	47.7	2237	US-08-354-973-1	Sequence 1, Appl
9	41	46.6	255	US-09-252-991A-19598	Sequence 19598, A
10	41	46.6	294	US-09-252-991A-21947	Sequence 21947, A
11	41	46.6	326	US-09-252-991A-20075	Sequence 20075, A
12	41	46.6	363	US-09-252-991A-25132	Sequence 25132, A
13	41	46.6	409	US-09-488-039A-14217	Sequence 14217, A
14	41	46.6	421	US-09-252-991A-32326	Sequence 32326, A
15	41	46.6	480	US-09-252-991A-32879	Sequence 32879, A
16	41	46.6	629	US-09-252-991A-22901	Sequence 22901, A
17	40.5	46.0	258	US-10-164-595-59	Sequence 59, Appl
18	40.5	46.0	274	US-09-976-594-417	Sequence 417, Appl
19	40.5	46.0	378	US-10-164-595-2	Sequence 2, Appl
20	40	45.5	239	US-09-252-991A-28985	Sequence 28985, A
21	40	45.5	255	US-09-342-084-11	Sequence 11, Appl
22	40	45.5	265	US-09-252-991A-19026	Sequence 19026, A
23	40	45.5	614	US-09-252-991A-29695	Sequence 29695, A
24	39.5	44.9	257	US-09-252-991A-22204	Sequence 22204, A
25	39.5	44.9	371	US-09-252-991A-29855	Sequence 29855, A
26	39	44.3	189	US-09-252-991A-24576	Sequence 24576, A
27	39	44.3	195	US-09-252-991A-30082	Sequence 30082, A

28	39	44.3	288	US-09-252-991A-27879	Sequence 27879, A
29	39	44.3	326	US-09-252-991A-17002	Sequence 17002, A
30	39	44.3	486	US-08-821-355A-8	Sequence 8, Appl
31	39	44.3	486	US-09-003-687A-8	Sequence 8, Appl
32	39	44.3	486	US-09-136-605-8	Sequence 8, Appl
33	39	44.3	491	US-09-489-039A-7836	Sequence 7836, A
34	39	44.3	154	US-09-252-991A-21434	Sequence 21434, A
35	38	43.2	157	US-09-252-991A-27836	Sequence 27836, A
36	38	43.2	274	US-09-252-991A-16696	Sequence 16696, A
37	38	43.2	279	US-09-252-991A-32606	Sequence 32606, A
38	38	43.2	329	US-09-252-991A-28088	Sequence 28088, A
39	38	43.2	379	US-09-252-991A-28428	Sequence 28428, A
40	38	43.2	380	US-09-252-991A-24774	Sequence 24774, A
41	38	43.2	392	US-09-252-991A-31291	Sequence 31291, A
42	38	43.2	447	US-09-189-637A-351	Sequence 351, Appl
43	38	43.2	508	US-08-818-024-3	Sequence 3, Appl
44	38	43.2	508	US-09-314-775A-3	Sequence 3, Appl
45	38	43.2	518	US-09-252-991A-25967	Sequence 25967, A

ALIGNMENTS

RESULT 1
US-09-513-057C-13
Sequence 13, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 653
TYPE: PRT
ORGANISM: Cardamine hirsputa
US-09-513-057C-13

Query Match 51.1%; Score 45; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

CY 1 SHOEYXGRSGRSGSGS 19
DB 573 SROVSTSSASGRGSGS 591
RESULT 2
US-09-984-890-4
Sequence 4, Application US/09984890
Patent No. 6492156
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CLO01306
CURRENT APPLICATION NUMBER: US/09/984,890
CURRENT FILING DATE: 2001-10-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 722
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match 50.0%; Score 44; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
DB 562 SPGHSGRGRGASGS 576

RESULT 3
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-890-2

Query Match 50.0%; Score 44; DB 4; Length 724;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
DB 564 SPGHSGRGRGASGS 578

RESULT 4
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match 50.0%; Score 44; DB 4; Length 745;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
DB 585 SPGHSGRGRGASGS 599

RESULT 5
US-09-342-084-6

; Sequence 6, Application US/09342084
; Patent No. 6251668
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Transcription Coactivators
; FILE REFERENCE: BB-1169-A
; CURRENT APPLICATION NUMBER: US/09/342,084
; CURRENT FILING DATE: 1999-06-28
; EARLIER APPLICATION NUMBER: 60/092,659
; EARLIER FILING DATE: July 13, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (128)
US-09-342-084-6

Query Match 47.7%; Score 42; DB 3; Length 169;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOEYGRGRGRGXSGS 19
| | | | |
DB 15 NKKSSGSSRGRTSPGS 33

RESULT 6
US-09-252-991A-24066
; Sequence 24066, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24066
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24066

Query Match 47.7%; Score 42; DB 4; Length 448;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 SHOEYGRGRGRGXSGS 18
| | | | |
DB 412 SHORGPAGRGARSRPPG 429

RESULT 7
US-08-817-832B-32
; Sequence 32, Application US/08817832B
; Patent No. 6579691
; GENERAL INFORMATION:
; APPLICANT: MANDELKOW, Eckhard, et al.
; TITLE OF INVENTION: NO. 6579691e1 Protein Kinase (NPK-110)
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower

```

CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,832B
FILING DATE: 28-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP95/04258
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94 11 7122.5
FILING DATE: 28-OCT-1994
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 722 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-832B-32.

Query Match 47.7%; Score 42; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0.

CY 5 STGGRGRGSGSGS 19
| : | | | | |
Db 562 SPGSHSGRRGPGSGS 576

RESULT 8
US-08-354-973-1
; Sequence 1, Application US/08354973
; Patent No. 564166
; GENERAL INFORMATION:
; APPLICANT: Vahlensteck, Hans-Friedrich
; APPLICANT: Hinnen, Albert
; TITLE OF INVENTION: Fungi Resistant to Scraphen A
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: Patent Dept., 520 White Plains Rd., POB 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/354,973
FILING DATE: 13-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: PF/5-19802/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2237 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
US-08-354-973-1

Query Match          47.7%; Score 42; DB 1; Length 2237;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches      9; Conservative      2; Mismatches      7; Indels      0; Gaps      0;

QY      2 HOESTXGRSGRSGXS GS 19
||| ||| ||| ||| |||
Db      1200 HQSSNGPAPDRGSSNS 1217

RESULT 9
US-09-252-991A-19598
Sequence 19598, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18598
LENGTH: 255
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19598

Query Match          46.6%; Score 41; DB 4; Length 255;
Best Local Similarity 58.3%; Pred. No. 33;
Matches      7; Conservative      3; Mismatches      2; Indels      0; Gaps      0;

QY      8 GRSGRSGXS GS 19
||| ||| ||| ||| |||
Db      64 GRDRGRAGAGA 75

RESULT 10
US-09-252-991A-21947
Sequence 21947, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21947
LENGTH: 294
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21947

Query Match          46.6%; Score 41; DB 4; Length 294;
Best Local Similarity 64.3%; Pred. No. 38;
Matches      9; Conservative      1; Mismatches      4; Indels      0; Gaps      0;

QY      5 STXGRSGRSGXS GS 18
||||| ||| ||| ||| |||

```

DB 22 SKSPSRGRSGPAG 35

RESULT 11

US-09-252-991A-20075
; Sequence 20075, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20075
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (15)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20075

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 326;
Best Local Similarity 64.3%; Pred. No. 42;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TXGRSGRSGXSGS 19

DB 272 TSPSRGRSGSGSS 285

RESULT 12

US-09-252-991A-25132
; Sequence 25132, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25132
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25132

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 363;
Best Local Similarity 53.3%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTYGRSGRSGS 15

DB 91 AHHSVGRGRWRG 105

RESULT 13

US-09-489-039A-14217
; Sequence 14217, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14217
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14217

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 409;
Best Local Similarity 47.4%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTYGRSGRSGXSGS 19

DB 374 SROASSPNEKSGRSGSGS 392

RESULT 14

US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 421;
Best Local Similarity 47.1%; Pred. No. 54;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 HOESTYGRSGRSGXSG 18

DB 13 HARTGTGTGRGRGRSG 29

RESULT 15

US-09-252-991A-22879
; Sequence 22879, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22879

Tue Sep 28 07:11:40 2004

us-09-308-150-9.open.ra1

Page 5

```
; LENGTH: 480
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22879
```

Query Match	46.6%	Score 41	DB 4	Length 480
Best Local Similarity	42.1%	Pred. No. 62		
Matches	8	Conservative	5	Mismatches 6
				Indels 0
				Gaps 0

```
Qy      1 SHQESTXGRSRGRSGXSGS 19
Db      52 SKDSSAGRARGSTGFPAS 70
```

```
Search completed: September 28, 2004, 06:26:46
Job time : 15.4479 secs
```

Blank

Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 14.4375 Seconds
(without alignment)
139.915 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HCHQESTXGRSGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	74.8	2248	2 A35938	profilaggrin - hum
2	79	68.7	416	2 A32947	flaggrin precursor
3	70	60.9	591	2 A45135	profilaggrin - hum
4	53	46.1	822	2 T51049	related to nucleol
5	51	44.3	506	1 W2W47	E2 protein - human
6	48.5	42.2	524	2 S28539	disintegrin-like m
7	48.5	42.2	670	2 I65967	disintegrin-like m
8	48	41.7	214	2 T45714	endochitinase-like
9	47	40.9	266	2 T14345	chitinase (EC 3.2.
10	47	40.9	266	2 T14348	probable chitinase
11	47	40.9	400	2 B48613	surface protein SU
12	47	40.9	730	2 S24376	2-aminobenzoyl-CoA
13	46.5	40.4	338	2 B75584	cytochrome-c perox
14	46	40.0	136	2 T35632	probable transpos
15	46	40.0	268	2 T14341	probable chitinase
16	46	40.0	849	2 A96592	hypothetical prote
17	45	39.1	202	2 T50635	hypothetical prote
18	45	39.1	268	2 T14344	chitinase (EC 3.2.
19	45	39.1	289	2 AD3394	lipoprotein releas
20	45	39.1	384	1 G70961	probable hexosyltr
21	44	38.3	440	2 T24232	hypothetical prote
22	44	38.3	825	2 UC4163	DNA-binding protei
23	44	38.3	836	2 G84727	probable DNA topoi
24	44	38.3	1215	2 I52882	autocatalytic huma
25	44	38.3	1804	2 T34518	nestin - golden ha
26	43	37.4	38	2 S74088	defensin - Mediter
27	43	37.4	334	2 G81744	hypothetical prote
28	43	37.4	1219	2 T61713	co-repressor prote
29	43	37.4	1669	1 CGWS4B	collagen alpha 1(I

30	42.5	37.0	299	2 T35765	hypothetical prote
31	42	36.5	123	2 T16234	hypothetical prote
32	42	36.5	323	2 D85074	hypothetical prote
33	42	36.5	424	2 T33552	hypothetical prote
34	42	36.5	519	2 T22091	hypothetical prote
35	42	36.5	682	2 T15092	hypothetical prote
36	42	36.5	888	2 S78288	preprotein translo
37	42	36.5	952	2 T18900	disintegrin and me
38	41.5	36.1	292	2 T03122	hypothetical prote
39	41.5	36.1	487	2 T06040	hypothetical prote
40	41	35.7	229	2 UC7219	nuclear protein SR
41	41	35.7	294	2 A96155	hypothetical prote
42	41	35.7	312	2 A31846	130K paracrystall
43	41	35.7	350	2 A85056	probable transpos
44	41	35.7	368	2 E87753	protein C43E11.5 (
45	41	35.7	368	2 T29779	hypothetical prote

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence__revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; MCBride, O.W.; Idler, W.W.; Markova, N.; Steiner, P.M.
Biochemistry 29, 9432-9440, 1990
A>Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GAN>
A:Cross-references: GB:002929
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
A:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: BP hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: flaggrin repeat
F:370-893/Region: flaggrin repeat
F:1074-1397/Region: flaggrin repeat
F:1573-1896/Region: flaggrin repeat

Query Match 74.8%; Score 86; DB 2; Length 2248;
Best local similarity 81.0%; Prod. No. 0.00016;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HCHQESTXGRSGRCGRSGS 21
Db 225 HOSHQESTXGRSGRCGRSGS 245

RESULT 2
A32947
flaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence__revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cammizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A>Title: Characterization of a cDNA clone encoding human flaggrin and localization of tr
A:Reference number: A32947; MUID:89256901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M24355; NID:g182604; PIND:AA52454.1; PTD:g182605
A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 68.7%; Score 79; DB 2; Length 416;
 Best Local Similarity 80.0%; Pred. No. 0.00045;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHQESTXGRSGRCGRSGS 21
 DB 6 QSHQESTRGRSAGRSGSGS 25

RESULT 3

A45135
 C:Species: Homo sapiens (man)
 C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135

R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nitrunksi, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992

A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717

A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-591 <PRE>
 A:Cross-references: GB:L01089; GB:M90957; NID:9190408; PIDN:AAA60177.1; PID:G553621

A:Note: sequence extracted from NCBI backbone (NCBI:P.118773)
 C:Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940

A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

F:49-81/Domain: calmodulin repeat homology <EP2>

Query Match 60.9%; Score 70; DB 2; Length 591;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 QCHQESTXGRSGRCGRSGS 21
 DB 448 QSHQESTRGRSAGRSGSGS 467

RESULT 4

T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12P1.10
 C:Species: Neurospora crassa

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049

R:Schulte, U.; Aign, V.; Hehseisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: Z25286

A:Accession: T51049

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AJ390091; GSPDB:GN00116; NCSP:B12P1.10

A:Experimental source: BAC clone B12P1; strain OR74A

C:Genetics:
 A:Gene: NCSP:B12P1.10

A:Map position: 6
 A:Introns: 80/2

Query Match 46.1%; Score 53; DB 2; Length 822;
 Best Local Similarity 58.8%; Pred. No. 4.6;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 QESTXGRSGRCGRSGS 21
 DB 414 RESACRTRGRGRCGT 430

RESULT 5

W2ML47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47

A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999

C:Accession: D35324
 R:Kiyono, T.; Adachi, A.; Ichibaashi, M.

Virology 177, 401-405, 1990

A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infer
 A:Reference number: A35324; MUID:90281611; PMID:2162112

A:Accession: D35324
 A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M32305; NID:9333062; PIDN:AAA46979.1; PID:9333067

C:Superfamily: Papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 44.3%; Score 51; DB 1; Length 506;
 Best Local Similarity 66.7%; Pred. No. 6.1;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 STXGRSGRCGRSGS 21
 DB 346 NTRGRGRCGRGAGS 360

RESULT 6

S38539
 disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 1 - human

N:Alternate names: ADAM1; MDC
 C:Species: Homo sapiens (man)

C:Date: 07-Oct-1994 #sequence_revision 17-Nov-1995 #text_change 26-May-2000
 C:Accession: S38539; I52965

R:Emi, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, J.; Ito, I.; Kasumi, F.; Nakamura
 Nature Genet. 5, 151-157, 1993

A:Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearrang
 A:Reference number: S38539; MUID:94073190; PMID:8252040

A:Accession: S38539
 A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <EMI>
 A:Cross-references: GB:D17390; NID:9452188; PIDN:BA404213.1; PID:d1004732; PID:g484255

R:Katagiri, T.; Harada, Y.; Emi, M.; Nakamura, Y.
 Cytogenet. Cell Genet. 68, 39-44, 1995

A:Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and
 A:Reference number: I52965; MUID:95044425; PMID:7956356

A:Accession: I52965

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 7-225, 'D', 227-524 <KAT>
 A:Cross-references: GB:D11872; NID:9505090; PIDN:BA06671.1; PID:d1007243; PID:g836684

C:Comment: For an alternative splice form, see PIR:I65967.
 C:Genetics:

A:Gene: GDB:MDC; ADAM1

A:Cross-references: GDB:230267; OMIM:155120

A:Map position: 17q21.3-17q21.3
 A:Introns: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341/

C:Superfamily: disintegrin homology
 C:Keywords: alternative splicing; hydrolase; metalloproteinase
 F:344-427/Domain: disintegrin homology <DIS>

Query Match 42.2%; Score 48.5; DB 2; Length 524;
 Best Local Similarity 45.0%; Pred. No. 14;
 Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHQESTXGRSGRCGRSGS 21

Db 468 CYEKLNVETGERSGCRKGS 487

RESULT 7

disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 2 - human
N/Alternate names: ADAM11; MDC
C/Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 17-Mar-2000 #text_change 11-Jan-2002
C/Accession: I65967; S38539
R/Katagiri, T.; Harada, Y.; Eml, M.; Nakamura, Y.
Cytogenet. Cell Genet. 68, 39-44, 1995
A/Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and
A/Reference number: I52965; MUID:95044425; PMID:7956356
A/Accession: I65967
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 7-670 <KAT>
A/Cross-references: GB:DJ3872; NID:9505090; PID:BA06670.1; PID:9836683
R/Eml, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, Y.; Ito, I.; Kasumi, F.; Nakamu
Nature Genet. 5, 151-157, 1993
A/Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearran
A/Reference number: S38539; MUID:94073190; PMID:8252040
A/Accession: S38539
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-225, 'N', 227-495, 'PQGRVWLPDLCQHLWSSSARPGGRHO' <EMT>
A/Cross-references: GB:DJ390; NID:9452188; PID:BA04213.1; PID:9484255
C/Comment: For an alternative splice form, see PIR:S38539.
C/Genetics:
A/Gene: GDB:MDC; ADAM11
A/Cross-references: GDB:230267; OMIM:155120
A/Map position: 17q21.3-17q21.3
A/Introns: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341
C/Superfamily: disintegrin homology
C/Keywords: alternative splicing; hydrolase; metalloproteinase
F/344-427/Domain: disintegrin homology <DIS>

Query Match 42.2%; Score 48.5; DB 2; Length 670;

Best Local Similarity 45.0%; Pred. No. 17;

Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHQE-STXGRSRGRCGRSGS 21

Db 468 CYEKLNVETGERSGCRKGS 487

RESULT 8

T45714

endochitinase-like protein - Arabidopsis thaliana

N/Alternate names: protein P1P2.90

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000

C/Accession: T45714

R/Choinsse, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, November 1999

A/Reference number: Z23010

A/Accession: T45714

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-214 <CHO>

A/Cross-references: EMBL:AL12955

A/Experimental source: cultivar Columbia; BAC clone F1P2

C/Genetics:

A/Map position: 3

A/Introns: 93/1

A/Note: P1P2.90

C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

Query Match 41.7%; Score 48; DB 2; Length 214;

Best Local Similarity 56.2%; Pred. No. 8.4;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 HQCHESTXGRSRGRC 16

Db 93 HFCYIEIKKRSRGRC 108

RESULT 9

T14345

chitinase (EC 3.2.1.14) EP3-3/E7, class IV - carrot

C/Species: Daucus carota (carrot)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000

C/Accession: T14345

R/Kragh, K.; De Vries, S.C.

submitted to the EMBL Data Library, March 1996

A/Reference number: Z17995

A/Accession: T14345

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-266 <KRA>

A/Cross-references: EMBL:U52847; NID:91549332; PID:G1549333

A/Experimental source: strain sg766 trophy

C/Genetics:

A/Gene: EP3

C/Function:

A/Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami

A/Pathway: polysaccharide degradation

C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

C/Keywords: glycosidase; hydrolase; polysaccharide degradation

F/21-56/Domain: hevein chitin-binding domain homology <HCB>

F/69-266/Domain: plant chitinase homology <PCH>

Query Match 40.9%; Score 47; DB 2; Length 266;

Best Local Similarity 50.0%; Pred. No. 14;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CHQE-STXGRSRGRC 16

Db 136 CHKEETNGRDXKC 149

RESULT 10

T14348

probable chitinase (EC 3.2.1.14) EP3B/E6, class IV - carrot

C/Species: Daucus carota (carrot)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000

C/Accession: T14348

R/Kragh, K.; De Vries, S.C.

submitted to the EMBL Data Library, March 1996

A/Reference number: Z17995

A/Accession: T14348

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-266 <KRA>

A/Cross-references: EMBL:U52848; NID:91549334; PID:G1549335

A/Experimental source: strain sg766 trophy

C/Genetics:

A/Note: EP3

C/Function:

A/Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami

A/Pathway: polysaccharide degradation

C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

C/Keywords: glycosidase; hydrolase; polysaccharide degradation

F/21-56/Domain: hevein chitin-binding domain homology <HCB>

F/69-266/Domain: plant chitinase homology <PCH>

Query Match 40.9%; Score 47; DB 2; Length 266;

Best Local Similarity 50.0%; Pred. No. 14;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CHQE-STXGRSRGRC 16

Db 136 CHKEETNGRDXKC 149

```

RESULT 11
C:Accession: B48613
C:Species: avian myeloblastosis virus
C:Date: 21-Jan-1994 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C:Accession: B48613
C:Title: V.; Boroughs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Perbal
Virology 195, 812-819, 1993
A:Title: Pathogenic potential of myeloblastosis-associated virus: implication of env pro
A:Reference number: A48613; MUID:93311743; PMID:8393249
A:Contents: MAV1(N)/2
A:Accession: B48613
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-400 <UD>
A:Note: sequence extracted from NCBI backbone (NCBI:P135488)
C:Superfamily: type C retrovirus env polypeptide

Query Match 40.9%; Score 47; DB 2; Length 400;
Best Local Similarity 38.1%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
DB 251 YNCSQEGCGCGCQYRCCKARS 271

RESULT 12
C:Accession: S24376
C:Species: Pseudomonas sp.
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 08-Oct-1999
C:Accession: S24376; S13857
C:Title: Schmidt, U.; Bokranz, M.; Fuchs, G.
Eur. J. Biochem. 207, 715-722, 1992
A:Title: Novel aerobic 2-aminobenzoate metabolism. Nucleotide sequence of the plasmid ca
P.
A:Reference number: S24374; MUID:92339462; PMID:1633822
A:Accession: S24376
A:Molecule type: DNA
A:Residues: 1-730 <ALT1>
A:Cross-references: EMBL:X66604; NID:G45865; PTDN:CAA47170.1; PID:G45868
A:Note: the authors translated the codon TGG for residue 680 as Asp, CGT for residue 703
Eur. J. Biochem. 194, 647-653, 1990
A:Title: Evidence that enzymes of a novel aerobic 2-amino-benzoate metabolism in dentitri
A:Reference number: S13857; MUID:91099342; PMID:2176602
A:Accession: S13857
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4, 'X', 6-20, 'D', 21-24 <ALT2>
C:Genetics:
A:Genome: Plasmid
C:Keywords: PAD; flavoprotein; homodimer; NAD; oxidoreductase
F:8-12/Region: PAD binding #status predicted
F:157-162/Region: NAD binding #status predicted

Query Match 40.9%; Score 47; DB 2; Length 730;
Best Local Similarity 52.6%; Pred. No. 31;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 19
DB 332 HQTAAESNIAWRCGRGA 350

RESULT 13
C:Accession: B75584
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Apr-2000

```

```

C:Accession: B75584
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Apr-2000
C:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75584
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <WHI>
A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12472.1; PID:9646076;
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0301
A:Map position: 2
C:Superfamily: Pseudomonas cytochrome-c peroxidase; Pseudomonas cytochrome-c peroxidase
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:72,75/Binding site: heme (Cys) (covalent) (low potential) #status predicted
F:76,273/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:216,219/Binding site: heme (Cys) (covalent) (high potential) #status predicted
F:220/Binding site: heme iron (His) (axial ligand) (high potential) #status predicted

Query Match 40.4%; Score 46.5; DB 2; Length 338;
Best Local Similarity 56.5%; Pred. No. 20;
Matches 13; Conservative 1; Mismatches 4; Indels 5; Gaps 2;

QY 3 CHQES---TXGR--SRGRCGRSG 20
DB 75 CHQSHAFTDGRATSLGTFGRSG 97

RESULT 14
C:Accession: T35632
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
C:Accession: T35632
C:Title: Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T35632
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-136 <SEB>
A:Cross-references: EMBL:AL079356; PIDN:CA845627.1; GSPDB:GN00070; SCQEDB:SC6G9.36C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCQEDB:SC6G9.36C
C:Superfamily: Streptomyces coelicolor probable transposase SC6G9.36C

Query Match 40.0%; Score 46; DB 2; Length 136;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 OCHOESTXGRSGRCGRSGS 19
DB 110 RAHQHAAARRGKRTGGA 127

RESULT 15
C:Accession: T14341
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14341
C:Title: Kraigh, K.; Toonen, M.A.J.; Bucherna, N.; Lo Schiavo, F.; Hendriks, T.; Weijer, E.A.; K
Plant Mol. Biol. 31, 631-645, 1996
A:Title: Characterization of carrot chitinases able to rescue the temperature-sensitive
A:Reference number: Z17992; MUID:96382431; PMID:8790295
A:Accession: T14341
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

A;Residues: 1-268 <KRA>
A;Cross-references: EMBL:U52845; NID:G1549328; PIDN:AAC49435.1; PID:G1549329
A;Experimental source: strain sg766 trophy
C;Genetics:
A;Note: EP3
C;Function:
A;Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami
A;Pathway: polysaccharide degradation
C;Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 40.0%; Score 46; DB 2; Length 268;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 CHOESTXGSRGRC 16
DB 138 CHOESTGDRHNYC 151

Search completed: September 28, 2004, 06:15:04
Job time : 16.4375 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.98438 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-10

Perfect score: 115
Sequence: 1 HQCHQESTXGRSGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	79	68.7	FILA_HUMAN	P20930 homo sapien
2	51	44.3	VE2_HPV47	P22420 human papill
3	49.5	43.0	AD11_MOUSE	Q9XIV4 mus musculu
4	48.5	42.2	AD11_HUMAN	Q75078 homo sapien
5	47.5	41.3	AKP1_RAT	O88884 r a kinase
6	46	40.0	PER2_AZOVI	P88802 azotobacter
7	45	39.1	Z205_HUMAN	O95201 homo sapien
8	44	38.3	5E5_RAT	O63003 ratu
9	44	38.3	M2GD_HUMAN	Q9U17 homo sapien
10	43	37.4	DEP1_MYTGA	P80571 mytilus gal
11	43	37.4	CA14_MOUSE	P02463 mus musculu
12	42.5	37.0	AD21_MOUSE	Q91476 mus musculu
13	42	36.5	YICG_CAEEL	Q8mu7 caenorhabdi
14	42	36.5	LSM4_CAEEL	Q19952 caenorhabdi
15	42	36.5	Y4CG_RHISN	P55389 rhizobium s
16	42	36.5	SECA_ODOSI	P49649 odontella s
17	42	36.5	ARRB_RAT	Q9e67 ratu
18	42	36.5	RM1_HUMAN	Q92445 homo sapien
19	42	36.5	PL15_CHICK	Q98917 gallu
20	41	35.7	MRK2_MOUSE	Q32522 mus musculu
21	41	35.7	SEAR_MYCLE	Q50612 mycobacteri
22	41	35.7	SEAR_MYCTU	Q50612 mycobacteri
23	41	35.7	8081_PHYB_OXYSA	P25764 oxyza sativ
24	40.5	35.2	RBPI_DROME	Q02427 drosophila
25	40.5	35.2	THIR_THERVO	Q974K6 thermoplas
26	40.5	35.2	AD20_HUMAN	Q43506 homo sapien
27	40.5	35.2	RRE1_HUMAN	Q92766 homo sapien
28	40.5	35.2	DEF1_MOUSE	Q91159 mus musculu
29	40	34.8	DEF1_ASCY	P80154 aescna cyz
30	40	34.8	PCRB_HALNI	Q9h4f8 halobacteri
31	40	34.8	PPNK_TREPA	O8345 treponema p
32	40	34.8	FCN1_MOUSE	Q70165 mus musculu
33	40	34.8	FCN1_RAT	Q9WES8 rattus norv

34	40	34.8	292	1	NOTC_RHIFR	P26508 rhizobium f
35	40	34.8	562	1	MTBE_MYCLE	Q9ccj1 mycobacteri
36	40	34.8	650	1	BM86_BOOMI	P20736 boophilus m
37	40	34.8	768	1	ITB8_RABIT	P26013 oryctolagus
38	40	34.8	816	1	AD15_RAT	Q9qyV0 r adam 15 p
39	40	34.8	857	1	AKP1_MOUSE	Q08715 mus musculu
40	40	34.8	1969	1	Z292_HUMAN	O60281 homo sapien
41	39	33.9	101	1	RS14_VIBCH	Q9kn27 vibrio chol
42	39	33.9	102	1	SNAS_MSAU	P81491 mesocricetu
43	39	33.9	106	1	COLA_HORSE	P02704 equus cabal
44	39	33.9	110	1	PER2_AQUAE	O66511 aquilex aeo
45	39	33.9	224	1	WFD5_HUMAN	O8rcv5 homo sapien

ALIGNMENTS

RESULT 1	FILA_HUMAN	STANDARD;	PRT;	416 AA.
ID	FILA_HUMAN			
AC	P20930	01-FEB-1991 (Rel. 17, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Flilaggrin precursor (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89296901; PubMed=2740331.			
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,			
RA	Caminizaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.,			
RT	"Characterization of a cDNA clone encoding human filaggrin and			
RT	localization of the gene to chromosome region 1q21."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).			
RN	[2]			
RP	CITRULLINATION.			
RX	MEDLINE=96374388; PubMed=8780679.			
RA	Senshu T., Kan S., Ogawa H., Manabe M., Asaga H.			
RT	"Preferential delamination of Keratin K1 and filaggrin during the			
RT	terminal differentiation of human epidermis."			
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).			
CC	-1- FUNCTION: Aggregates keratin intermediate filaments and promotes			
CC	disulfide-bond formation among the intermediate filaments during			
CC	terminal differentiation of mammalian epidermis.			
CC	-1- PTM: Filaggrin is initially synthesized as a large, insoluble,			
CC	highly phosphorylated precursor containing many tandem copies of			
CC	324 AA, which are not separated by "large linker". The precursor			
CC	is deposited as keratohyalin granules. During terminal			
CC	differentiation it is dephosphorylated and proteolytically			
CC	cleaved.			
CC	-1- PTM: Undergoes delamination of some arginine residues			
CC	(citrullination).			
CC	-----			
CC	THE SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL, outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M24355; AA52454.1; -			
CC	PIR; A32947; A32947.			
CC	Genew: HGNC:3748; FLG.			
CC	MIM: 135940;			
CC	GO: 0005882; C: intermediate filament; NAS.			
CC	GO: 0005198; F: structural molecule activity; NAS.			
CC	GO: 0007275; P: development; NAS.			
CC	InterPro: IPR003103; Flilaggrin.			

DR Pfam: PF03515; Filaggrin; 2.
 DR PRINTS: PR00487; FILAGRIN.
 KW Phosphorylation; Cytullination; Developmental protein.
 FT NON_TER 1

SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 68.7%; Score 79; DB 1; Length 416;
 Best Local Similarity 80.0%; Pred. No. 0.00012;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHQSTXGRSGRGRSGS 21
 Db 6 QSHQSTXGRSGRGRSGS 25

RESULT 2

VE2_HPV47 STANDARD; PRT; 506 AA.
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2.
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 CX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9028161; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).

CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2B RESPONSE ELEMENT (5'-ACCCNNNNNGGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2B'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY SERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33305; AAA46979.1; -
 DR PIR: D35324; W2WL47.
 DR HSSP: P03122; 2BOP.
 DR InterPro: IPR000427; E2_C.
 DR InterPro: IPR001866; E2_N.
 DR InterPro: IPR009021; Viral_DNA_bd.
 DR Pfam: PF00511; E2_C; 1.
 DR Pfam: PF00508; E2_N; 1.
 DR ProDom: PD000672; E2_C; 1.
 DR ProDom: PD000678; E2_N; 1.
 DR Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF8725065 CR654;

Query Match 44.3%; Score 51; DB 1; Length 506;
 Best Local Similarity 66.7%; Pred. No. 2.1;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 STXGRSGRGRSGS 21

Db 346 NTRGRGRGRGRS 360

RESULT 3

AD11_MOUSE STANDARD; PRT; 773 AA.

AC Q9RIV4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ADAM 11 precursor (A disintegrin-like and metalloproteinase domain 11)
 DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)
 GN (MDC).
 GN ADAM11 OR MDC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99365303; PubMed=10433968;
 RA Sagane K., Yamazaki K., Mizui Y., Tanaka I.;
 RT "Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23.";
 RL Gene 236:79-86(1999).

CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non
 CC catalytic metalloproteinase-like protein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in the brain. Weakly
 CC detected in the heart, liver and testis.
 CC -1- DOMAIN: A conserved motif [AVN(E/D)CD] within the disintegrin-like
 CC domain could be involved in the binding to the integrin receptor.
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 disintegrin domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB009676; BA83384.1; -
 DR HSSP: P18619; 1EVL.
 DR MEROPS: M12.976; -
 DR MGD: MGI:1096667; Adam11.
 DR InterPro: IPR006586; ADAM cysteine.
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR006209; EGF like.
 DR InterPro: IPR006025; Pept_M_2n_BS.
 DR InterPro: IPR001590; Peptidase_M12B.
 DR InterPro: IPR002870; Peptidase_M12B_N.
 DR Pfam: PF00200; disintegrin; 1.
 DR Pfam: PF01562; Pep_M12B_propep; 1.
 DR Pfam: PF01421; Repolysein; 1.
 DR PRINTS: PR00289; DISINTEGRIN.
 DR ProDom: PD000664; Disintegrin; 1.
 DR SMART: SMO0608; ACR; 1.
 DR SMART: SMO050; DISIN; 1.
 DR PROSITE: PS50215; ADAM_MEROP; 1.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS50214; DISINTEGRIN_2; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_3; FALSE_NEG.
 DR PROSITE: PS50026; EGF_3; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain.
 FT SIGNAL 1 24 POTENTIAL.

```

FT PROPEP 25 229 BY SIMILARITY.
FT CHAIN 230 773 ADAM 11.
FT DOMAIN 230 738 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 739 759 POTENTIAL.
FT DOMAIN 760 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 230 773 METALLOPROTEASE-LIKE.
FT DOMAIN 448 535 DISINTEGRIN-LIKE.
FT DOMAIN 536 680 CYS-RICH.
FT DOMAIN 681 713 EGF-LIKE.
FT DISULFID 353 437 BY SIMILARITY.
FT DISULFID 507 520 POTENTIAL.
FT DISULFID 681 696 BY SIMILARITY.
FT DISULFID 690 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 773 AA; 84134 MW; 9A8A125FB3F743DD CRC64;

Query Match 43.0%; Score 49.5; DB 1; Length 773;
Best Local Similarity 45.0%; Pred. No. 5.4;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

3 CHO2-STXGRSGRCGRSGS 21
571 CYEKLWVETGRNGCRKGS 590

RESULT 4
AD1_HUMAN STANDARD; PRT: 769 AA.
ID 075078; Q14808; Q14809; Q14810.
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAM 11 precursor (A disintegrin and metalloproteinase domain 11)
DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)
DE (MDC)
GN ADAM11 OR MDC.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=98359734; PubMed=693107;
RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.,
RT "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2
and MDC3: novel human cellular disintegrins highly expressed in the
brain.";
RL Biochem. J. 334:93-98(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Cerebellum;
RX MEDLINE=94073190; PubMed=8252040;
RA Emti M., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I.,
RT Kasumi F., Nakamura Y.;
RA "A novel metalloproteinase/disintegrin-like gene at 17q21.3 is
somatic ally rearranged in two primary breast cancers.";
RL Nat. Genet. 5:151-157(1993).
RN [3]
RP SEQUENCE OF 106-769 FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=Brain, Breast, Ovary, and Testis;
RX MEDLINE=95044425; PubMed=7956356;
RA Katagiri T., Harada Y., Emti M., Nakamura Y.;
RT "Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron
organization and alternative splicing.";
RL Cytogenet. Cell Genet. 68:39-44(1995).
CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non
catalytic metalloproteinase-like protein.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

```

```

CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long; Synonyms=MDC-769;
CC Name=Short; Synonyms=MDC-524;
CC IsoId=O75078-1; Sequence=Displayed;
CC IsoId=O75078-2; Sequence=VSP_005472, VSP_005473, VSP_005474,
VSP_005475;
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain. Slightly
detected or not at all in other tissues.
CC -1- DOMAIN: A conserved motif [AVNIE/D/CD] within the disintegrin-like
domain could be involved in the binding to the integrin receptor.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -1- DISEASE: Rearrangements occur in breast and ovarian cancers which
involve multiple exons and disrupt the coding region.
CC -1- SIMILARITY: Belongs to peptidase family M12b.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 disintegrin domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009675; BAA32352.1; -.
DR EMBL; D17390; BAA04213.1; -.
DR EMBL; D31872; BAA06670.1; -.
DR EMBL; D31872; BAA06671.1; -.
DR PIR; I65967; I65967.
DR HSSP; P18619; 1FVL.
DR MEROPS; M12.976; -.
DR GeneW; HGNC:189; ADAM11.
DR MIM; 155120; -.
DR GO; GO:0005886; C:Plasma membrane; TAS.
DR GO; GO:0005178; F:integrin binding; TAS.
DR GO; GO:0008237; F:metalloproteinase activity; TAS.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006025; Pept_M_2n_BS.
DR InterPro; IPR001590; Peptidase_M12b.
DR InterPro; IPR002870; Peptidase_M12b_N.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; pep_M12b_propep; 1.
DR Pfam; PF01421; Reptolysin_1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD00664; Disintegrin; 1.
DR SMART; SM00608; ACR1; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEROP; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR Signal; Glycoprotein; Transmembrane; EGF-like domain;
Alternative splicing;
KM SIGNAL 1 23
FT PROPEP 24 225 BY SIMILARITY.
FT CHAIN 226 769 ADAM 11.
FT DOMAIN 226 734 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 735 755 POTENTIAL.
FT DOMAIN 756 769 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 226 438 METALLOPROTEASE-LIKE.
FT DOMAIN 444 531 DISINTEGRIN-LIKE.
FT DOMAIN 532 676 CYS-RICH.
FT DOMAIN 677 709 EGF-LIKE.
FT DISULFID 349 433 BY SIMILARITY.

```

```

FT DISULFID 503 516 POTENTIAL.
FT DISULFID 677 692 BY SIMILARITY.
FT DISULFID 686 698 BY SIMILARITY.
FT DISULFID 700 709 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASAPLIC 1 99 Missing (in isoform Short).
FT VASAPLIC 100 104 /FtId=VSP_005472.
FT VASAPLIC 595 623 /FtId=VSP_005473.
FT VASAPLIC 595 623 DVLGFLCYNIGAPFLDVLGDISVT -> POCGRAMW
FT VASAPLIC 595 623 LPLCGLHMSARGPCGRHQ (in isoform Short).
FT VASAPLIC 624 769 /FtId=VSP_005474.
FT VASAPLIC 624 769 Missing (in isoform Short).
FT VASAPLIC 106 106 /FtId=VSP_005475.
FT VASAPLIC 325 325 Q -> H (IN REF. 1).
FT VASAPLIC 325 325 D -> N (IN REF. 2).
SQ SEQUENCE 769 AA; 83409 MW; 59B9C467B6DDF86E CRC64;

Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred. No. 7.6;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHOESTXGRSGRGRSGS 21
Db 567 CYEXLNEGTGRSGCRKGS 586

RESULT 5
AKP1_RAT STANDARD; PRT; 854 AA.
ID AKP1_RAT
AC O88884; O88884;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE A kinase anchor protein 1, mitochondrial precursor (Protein kinase A
DE anchoring protein 1) (PKA1) (A-kinase anchor protein 121 kDa) (AKAP
DE 121) (Dual specificity A-kinase anchoring protein 1) (D-AKAP-1)
DE (Spermatid A-kinase anchor protein 84) (S-AKAP84).
GN AKAP1 OR AKAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10116;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Thyroid;
RX MEDLINE=9389771; PubMed=9722570;
RA Feliciello A.; Rubin C.S.; Avvedimento E.V.; Gottesman M.E.;
RT "Expression of a kinase anchor protein 121 is regulated by hormones in
RT thyroid and testicular germ cells.";
RL J. Biol. Chem. 273:23361-23366(1998).
CC -1- FUNCTION: BINDS TO TYPE I AND II REGULATORY SUBUNITS OF PROTEIN
CC KINASE A AND ANCHORS THEM TO THE CYTOPLASMIC FACE OF THE
CC MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Mitochondrial outer membrane (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=AKAP121;
CC IsoId=O88884-1; Sequence=Displayed;
CC Name=2; Synonyms=S-AKAP84;
CC IsoId=O88884-2; Sequence=VSP_002854, VSP_002855;
CC -1- TISSUE SPECIFICITY: testis specific.
CC -1- INDUCTION: BY THYROID STIMULATING HORMONE (TSH) AND CAMP OR CAMP-
CC ANALOG.
CC -1- DOMAIN: RII-alpha binding site, predicted to form an amphipathic
CC helix, could participate in protein-protein interactions with a
CC complementary surface on the R-subunit dimer.
CC -1- SIMILARITY: Contains 1 KH domain.
CC -1- SIMILARITY: Contains 1 Tudor domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF068202; AAC61775.1; -
DR EMBL; AF092523; AAC61775.1; -
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR004191; Maternal_tudor.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00013; KH; 1.
DR Pfam; PF00567; TUDOR; 1.
DR SMART; SM00322; TUDOR; 1.
DR SMART; SM00333; TUDOR; 1.
DR PROSITE; PS50084; KH TYPE 1; 1.
DR PROSITE; PS50304; TUDOR; 1.
DR RNA-binding; Mitochondrion; Alternative splicing; Outer membrane;
DR Transmembrane; Transic peptide.
KW TRANSMIT 1 29 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 30 854
FT DOMAIN 303 316 A KINASE ANCHOR PROTEIN 1.
FT DOMAIN 558 622 PKA-RII SUBUNIT BINDING DOMAIN.
FT DOMAIN 709 768 KH.
FT VASAPLIC 523 544 TUDOR.
FT GSDNSNDSDVSCGLTKPSP -> VAPQGRHGNGG
FT CAGFFEC (in isoform 2).
FT VASAPLIC 545 854 /FtId=VSP_002854.
FT VASAPLIC 545 854 Missing (in isoform 2).
FT CONFLICT 311 311 I -> T (IN REF. 1; AAC61775).
FT CONFLICT 439 439 V -> A (IN REF. 1; AAC61775).
SQ SEQUENCE 854 AA; 91447 MW; 6EFBA30F8801A06E CRC64;

Query Match 41.3%; Score 47.5; DB 1; Length 854;
Best Local Similarity 47.8%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

QY 3 CHOESTXGRSGRGRSGS 20
Db 160 CKQSAJGRTPGRGWSQCAASG 182

RESULT 6
FER2_AZOVI STANDARD; PRT; 106 AA.
ID FER2_AZOVI
AC P82802.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ferredoxin, 2Fe-2s (Shetna protein I) (Iron-sulfur protein 2)
DE (2FeAVFd1).
OS Azotobacter vinelandii.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC Pseudomonadaceae; Azotobacter.
OX NCBI_taxid=354;
RN (1)
RP SEQUENCE, SUBUNIT, AND MASS SPECTROMETRY.
RC STRAIN=DJ;
RX MEDLINE=89367931; PubMed=10439076;
RA Chatelet C.; Meyer J.;
RT "The (2Fe-2S) protein I (Shetna protein I) from Azotobacter
RT vinelandii is homologous to the (2Fe-2S) ferredoxin from Clostridium
RT pasteurianum.";
RL J. Biol. Inorg. Chem. 4:311-317(1999).
CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS. MIGHT BE
CC INVOLVED IN NITROGEN FIXATION.
CC -1- COFACTOR: Binds 1 2Fe-2S cluster.
CC -1- SUBUNIT: Homodimer.

```



```
CC -1- MASS SPECTROMETRY: MW=22876; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE 2FE2S SHELTHNA-TYPE FERREDOXIN FAMILY.
DR HSSP; 066511; 1F37.
KW Electron transport; Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KW Nitrogen fixation.
FT METAL 10 10 IRON-SULFUR (2FE-2S)
FT METAL 23 23 IRON-SULFUR (2FE-2S)
FT METAL 55 55 IRON-SULFUR (2FE-2S)
FT METAL 59 59 IRON-SULFUR (2FE-2S)
SQ SEQUENCE 106 AA; 11264 MW; 803BB1158A96C6B7 CRC64;

Query Match 40.0%; Score 46; DB 1; Length 106;
Best Local Similarity 42.1%; Pred. No. 2.6;
Matches 8; Conservative 1; Mismatches 10; Gaps 0;

OY 3 CHOESTXGRSGRCGRSGS 21
DB 10 CAQNRPAHPRGSCGAKGA 28

RESULT 7
ID 2205 HUMAN STANDARD; PRT; 504 AA.
AC 095201;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 205 (Zinc finger protein 210).
GN ZNF205 OR ZNF210.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=9900537; Pubmed=9787081;
RA Deng Z., Centola M., Chen X., Sood R., Vedula A., Fischel-Ghodsian N.,
RA Kastner D.L.;
RT "Identification of two Kruppel-related zinc finger genes (ZNF200 and
RT ZNF210) from human chromosome 16p13.3."
RL Genomics 53:97-103(1998).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION REPRESSOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, SKELETAL MUSCLE, PANCREAS
CC AND BRAIN. WEAKLY EXPRESSED IN PLACENTA, LUNG, LIVER, KIDNEY AND
CC THYMUS.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF060865; AAC70007.1; -.
DR HSSP; P08046; 1A1G.
DR Genew; HGNC:12996; ZNF205.
DR MIM; 603436; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003677; F:DNA binding; NAS.
DR GO; GO:0008270; F:zinc ion binding; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 8.
DR Prodom; PD000003; Znf_C2H2; 8.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 8.
```

```
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repressor; Repeat.
FT DOMAIN 74 143 KRAB.
FT ZN_FING 258 280 C2H2-TYPE.
FT ZN_FING 286 308 C2H2-TYPE.
FT ZN_FING 314 336 C2H2-TYPE.
FT ZN_FING 342 364 C2H2-TYPE.
FT ZN_FING 370 392 C2H2-TYPE.
FT ZN_FING 398 420 C2H2-TYPE.
FT ZN_FING 426 448 C2H2-TYPE.
FT ZN_FING 454 476 C2H2-TYPE.
SQ SEQUENCE 504 AA; 55312 MW; 33C0430D748ABF8C CRC64;

Query Match 39.1%; Score 45; DB 1; Length 504;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 2 QCHOESTXGRSGRCGRSG 20
DB 184 QCAQEPACGRSGPAPDGS 202

RESULT 8
ID 585 RAT STANDARD; PRT; 825 AA.
AC 063003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 585 antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=wislar; TISSUE=Brain;
RX MEDLINE=9601519; Pubmed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Ujemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein 585 in the nervous system."
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D37934; BAA07153.1; -.
DR PIR; JC4163; JC4163.
DR DNA-binding; Nuclear protein; Antigen.
KW SEQUENCE 825 AA; 86831 MW; AF67FE2PDS55BDF CRC64;

Query Match 38.3%; Score 44; DB 1; Length 825;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 10 GRSRGRCGRSGS 21
DB 589 GRGRGRGRGRGS 600

RESULT 9
M2GD_HUMAN
```

```

ID      M2GD_HUMAN          STANDARD;          PRT;           866 AA.
AC      O9uni17.
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Dimethylglycine dehydrogenase, mitochondrial precursor (EC 1.5.99.2)
DE      (ME2GSDH).
GN      DMGDH.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_taxid=9606;
PI      [1]
RP      SEQUENCE FROM N.A., AND VARIANTS PRO-279; ALA-530 AND SER-546.
RX      MEDLINE=20232164; PubMed=10761772;
RA      Binzak B.A., Vockley J.G., Jenkins R.B., Vockley J.;
RT      "Structure and analysis of the human dimethylglycine dehydrogenase
RT      gene.";
RL      Mol. Genet. Metab. 69:181-187(2000).
RN      [2]
RN      SEQUENCE FROM N.A., AND VARIANT DMGDHD ARG-109.
RX      MEDLINE=21152272; PubMed=11231903;
RA      Binzak B.A., Wevers R.A., Moelenaar S.H., Lee Y.-M., Huw W.-L.,
RA      Poggi-Bach J., Engelse U.F.H., Hoard H.M., Vockley J.G., Vockley J.;
RT      "Cloning of dimethylglycine dehydrogenase and a new human inborn error
RT      of metabolism, dimethylglycine dehydrogenase deficiency.";
RL      Am. J. Hum. Genet. 68:833-847(2001).
RN      [3]
RE      DISEASE.
RP      MEDLINE=99203382; PubMed=10102904;
RA      Moelenaar S.H., Poggi-Bach J., Engelse U.F.H., Corstjansen J.M.B.,
RA      Heerschap A., de Jong J.G.N., Binzak B.A., Vockley J., Wevers R.A.;
RT      "Defect in dimethylglycine dehydrogenase, a new inborn error of
RT      metabolism: NMR spectroscopy study.";
RL      Clin. Chem. 45:459-464(1999).
CC      -1- CATALYTIC ACTIVITY: N,N-dimethylglycine + acceptor + H(2)O =
CC      sarcosine + formaldehyde + reduced acceptor.
CC      -1- COFACTOR: FAD; covalently linked.
CC      -1- PATHWAY: Choline metabolism.
CC      -1- SUBUNIT: Monomer.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial.
CC      -1- DISEASES: Defects in DMGDH are the cause of DMGDH deficiency
CC      (DMGDHD) [MIM:605850]. DMGDHD is a disorder characterized by fish
CC      odor, muscle fatigue with increased serum creatine kinase.
CC      Biochemically it is characterized by an increase of N,N-
CC      dimethylglycine (DMG) in serum and urine.
CC      -1- SIMILARITY: Belongs to the gcvt family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF111858; AAF21941.1; -.
DR      MIM; 605849; -.
KW      MIM; 605850; -.
DR      GO; GO:0005739; C:mitchondrion; TAS.
DR      GO; GO:0008246; P:electon transfer flavoprotein; TAS.
DR      InterPro; IPR0060272; Fcd oxred.
DR      InterPro; IPR006222; GCV_T.
DR      InterPro; IPR000205; NAD_BS.
DR      Pfam; PF01266; DAO; 1.
DR      Pfam; PF01571; GCV_T; 1.
KW      Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide;
KW      Disease mutation; Polymorphism.
FT      TRANSIT             1       50      MITOCHONDRION (POTENTIAL).
FT      CHAIN               51      866     DIMETHYLGLYCINE DEHYDROGENASE.
FT      BINDING            91      91       FAD (COVALENT) (BY SIMILARITY).
FT      VARIANT            109      109     H -> R (in DMGDHD).
FT                                     /FTId=VAR_011505.

```

```

FT    VARIANT      279      279      S -> P (1n dBSNP:552964) .
FT    VARIANT      530      530      G-> A (1n dBSNP:1805073) .
FT    VARIANT      646      646      P-> S (1n dBSNP:1805074) .
FT    SEQUENCE     866 AA; 96806 MW; 546B1FC1656AC04 CRC64;
SQ
Query Match      38.3%; Score 44; DB 1; Length 866;
Best Local Similarity 44.4%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0
QY      3 CHOESTXGRSGRCGRSG 20
Db      17 CPGGSPGRFRCVCCREG 34
RESULT 10
DEF1_MYTGA
ID    DEF1_MYTGA      STANDARD; PRT; 38 AA.
AC    P80571;
DC    01-OCT-1996 (rel. 34, Created)
DT    01-OCT-1996 (rel. 34, Last sequence update)
DT    10-OCT-2003 (rel. 42, Last annotation update)
DE    Defensin MGD-1.
GN
OS    Mytilus galloprovincialis (Mediterranean mussel).
OC    Eukaryota, Metazoa, Mollusca, Bivalvia, Pteriomorpha, Mytiloidea;
OC    Mytiloidea, Mytilidae; Mytilus.
NCBI_NCB1_TaxID=29158;
RN
RP    SEQUENCE.
RC    TISSUE=Hemolymph; PubMed=8925841;
RX    MEDLINE=9702539; PubMed=8925841;
RA    Hubert F., Noeël T., Roch P.;
RT    "A member of the arthropod defensin family from edible Mediterranean
RL    mussels (Mytilus galloprovincialis).";
RL    Eur. J. Biochem. 240:302-306 (1996).
[2]
RN
RP    ERRATUM.
RA    Hubert F., Noeël T., Roch P.;
RL    Eur. J. Biochem. 240:815-815 (1996).
[3]
RP    STRUCTURE BY NMR.
RX    MEDLINE=20541401; PubMed=11087396;
RA    Yang Y.S., Mlita G., Chavanieu A., Calas B., Sanchez J.F., Roch P.,
RA    Ameelaes A.;
RT    "Solution structure and activity of the synthetic four-disulfide bond
RL    Mediterranean mussel defensin (MGD-1).";
RT    Biochemistry 39:14436-14447 (2000).
CC    -1- SUBCELLULAR LOCATION: Secreted.
CC    -1- SIMILARITY: Belongs to the arthropod defensin family. Subfamily 2.
PIR, S74088; S74088.
PDB, 1FVN; 20-DEC-00.
DR    InterPro; IPR001542; Defensin_arnod.
DR    Pfam; PF01097; Arthro_defensin; 1.
DR    PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW    Antibiotic; 3D-structure.
FT    DISULFID 4 25
FT    DISULFID 10 33
FT    DISULFID 14 35
FT    DISULFID 21 38
SQ    SEQUENCE 38 AA; 4275 MW; 6F1CCDCCF0E69E76 CRC64;
Query Match      37.4%; Score 43; DB 1; Length 38;
Best Local Similarity 47.4%; Pred. No. 2; 7'
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;
QY      1 HCHQD--ESTXGRSGRCG 17
Db      8 YQCHRNCKSPGRGCGYC 26

```

```

RESULT 11
CAL4 MOUSE STANDARD; PRT; 1669 AA.
ID _CAL4 MOUSE
AC P02463;
RT 21-ntf-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Collagen alpha 1(IV) chain precursor.
GN COL4A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89197932; PubMed=2703490;
RA Muthukumaran G., Blumberg B., Kurkinen M.;
RT "The complete primary structure for the alpha 1-chain of mouse
RT collagen IV. Differential evolution of collagen IV domains.";
RT J. Biol. Chem. 264:6310-6317(1989).
RN [2]
RP SEQUENCE OF 1-1154 FROM N.A.
RX MEDLINE=88112221; PubMed=3338568;
RA Wood L., Theriault N., Vogel G.;
RT "cDNA clones completing the nucleotide and derived amino acid
RT sequence of the alpha 1 chain of basement membrane (type IV) collagen
RT from mouse.";
RT FEBS Lett. 227:5-8(1988).
RN [3]
RP SEQUENCE OF 1149-1424 FROM N.A.
RX MEDLINE=86301886; PubMed=3755692;
RA Nath P., Laurent M., Horn E., Sobel M.E., Zon G., Vogel G.;
RT "Isolation of an alpha 1 type-IV collagen cDNA clone using a
RT synthetic oligodeoxynucleotide.";
RT Gene 43:301-304(1986).
RN [4]
RP SEQUENCE OF 1276-1669 FROM N.A.
RX MEDLINE=85127033; PubMed=2578961;
RA Oberhaeumer I., Laurent M., Schwarz U., Sakurai Y., Yamada Y.,
RA Vogel G., Voss T., Siebold B., Glanville R.W., Kuhn K.;
RT "Amino acid sequence of the non-collagenous globular domain (NC1) of
RT the alpha 1(IV) chain of basement membrane collagen as derived from
RT complementary DNA.";
RT Eur. J. Biochem. 147:217-224(1985).
RN [5]
RP SEQUENCE OF 1441-1669 FROM N.A.
RX MEDLINE=87250460; PubMed=3597383;
RA Kurkinen M., Condon M.R., Blumberg B., Barlow D., Quinones S.,
RA Saus J., Philatjanemi T.;
RT "Extensive homology between the carboxyl-terminal peptides of mouse
RT alpha 1(IV) and alpha 2(IV) collagen.";
RT J. Biol. Chem. 263:8496-8499(1987).
RN [6]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=86196099; PubMed=3009468;
RA Sakurai Y., Sullivan M., Yamada Y.;
RT "Alpha 1 type IV collagen gene evolved differently from fibrillar
RT collagen genes.";
RT J. Biol. Chem. 261:6654-6657(1986).
RN [7]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89066738; PubMed=3198626;
RA Kayes P., Wood L., Theriault N., Kurkinen M., Vogel G.;
RT "Head-to-head arrangement of murine type IV collagen genes.";
RT J. Biol. Chem. 263:19274-19277(1988).
RN [8]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89071759; PubMed=3200851;
RA Burdalo P.D., Martin G.R., Yamada Y.;
RT "Alpha 1(IV) and alpha 2(IV) collagen genes are regulated by a
RT bidirectional promoter and a shared enhancer.";
RT Proc. Natl. Acad. Sci. U.S.A. 85:9679-9682(1988).
RN [9]

```

```

RP SEQUENCE OF 1-129 FROM N.A.
RX MEDLINE=88243724; PubMed=379041;
RA Killen P.D., Burdalo P., Sakurai Y., Yamada Y.;
RT "Structure of the amino-terminal portion of the murine alpha 1(IV)
RT collagen chain and the corresponding region of the gene.";
RT J. Biol. Chem. 263:8706-8709(1988).
RN [1]
RP FUNCTION: Type IV collagen is the major structural component of
RN glomerular basement membranes (GBM), forming a 'chicken-wire'
RN network together with laminins, proteoglycans and enactin/
RN nidogen.
CC [1] SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)-
CC alpha 6(IV), each of which can form a triple helix structure with
CC 2 other chains to generate type IV collagen network.
CC [1] DOMAIN: Alpha chains of type IV collagen have a noncollagenous
CC domain (NC1) at their C-terminus, frequent interruptions of the G-
CC X-Y repeats in the long central triple-helical domain (which may
CC cause flexibility in the triple helix), and a short N-terminal
CC triple-helical 7S domain.
CC [1] PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC [1] PTM: Type IV collagens contain numerous cysteine residues which
CC are involved in inter- and intramolecular disulfide bonding. 12 of
CC these, located in the NC1 domain, are conserved in all known type
CC IV collagens.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03758; AAA37439.1; -
DR EMBL; M23333; AAA51625.1; -
DR EMBL; U04694; AAA50292.1; -
DR EMBL; U06777; CAA29946.1; -
DR EMBL; X02201; CAA26132.1; -
DR EMBL; M15832; AAA37340.1; -
DR EMBL; M14042; AAA37342.1; -
DR EMBL; M12879; AAA37343.1; -
DR EMBL; M13024; -; NOT_ANNOTATED_CDS.
DR EMBL; M13025; -; NOT_ANNOTATED_CDS.
DR EMBL; M13026; AAA37344.1; -
DR EMBL; M13027; AAA37345.1; -
DR EMBL; M13043; AAA37346.1; -
DR EMBL; J04448; AAA37437.1; -
DR PIR; A33585; CGMS4B.
DR MGB; MGI:88454; COL4A1.
DR GO; GO:0005604; C:Basement membrane; IDA.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001442; Procollagen4_C.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 23.
DR ProDom; PD000007; C1g_helix; 6.
DR ProDom; PD003923; ProcollagenC4; 1.
DR SMART; SM00111; C4; 2.
DR ExTracellul; Extracellular matrix; Connective tissue; Basement membrane;
KW Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
FT SIGNAL; 1..27
FT PROPEP; 28..162
FT CHAIN; 173..1440
FT DOMAIN; 173..1440
FT TRIPLE-HELICAL REGION; 1441..1669
FT NONHELICAL REGION (NC1); 1441..1669
FT DISULFD; 1460..1551
FT DISULFD; 1493..1548
FT DISULFD; 1505..1511
FT DISULFD; 1570..1665
FT DISULFD; 1604..1662
FT DISULFD; 1616..1622
FT CARBOHYD; 126..126
FT CONFLICT; 26..26
FT AMINO-TERMINAL PROPEPTIDE (7S DOMAIN).
FT TRIPLE-HELICAL REGION.
FT NONHELICAL REGION (NC1).
FT OR 1548 (BY SIMILARITY).
FT OR 1551 (BY SIMILARITY).
FT BY SIMILARITY.
FT OR 1662 (BY SIMILARITY).
FT OR 1665 (BY SIMILARITY).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT A -> P (IN REF. 2).

```

```

FT CONFLICT 186 186 S -> L (IN REF. 2).
FT CONFLICT 319 319 Q -> S (IN REF. 2).
FT CONFLICT 369 369 L -> L (IN REF. 2).
FT CONFLICT 403 403 L -> F (IN REF. 2).
FT CONFLICT 481 481 P -> L (IN REF. 2).
FT CONFLICT 493 493 Q -> H (IN REF. 2).
FT CONFLICT 712 712 S -> I (IN REF. 2).
FT CONFLICT 813 813 E -> Q (IN REF. 2).
FT CONFLICT 982 982 Q -> H (IN REF. 2).
FT CONFLICT 1397 1397 V -> S (IN REF. 3).
SQ SEQUENCE 1669 AA; 160680 MW; 42916591E52058E9 CRC64;

Query Match 37.4%; Score 43; DB 1; Length 1669;
Best Local Similarity 41.2%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 HQESTXGRSGRCGRSG 20
DB 19 HERSRMAAKGDCGSG 35

RESULT 12
AD21_MOUSE STANDARD; PRT; 729 AA.
ID AD21_MOUSE
AC Q90T76;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAM 21 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE domain 21) (ADAM 31).
GN ADAM21 OR ADAM31.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Bone marrow;
RX MEDLINE=20289099; PubMed=10830287;
RA Liu L., Smith J.W.;
RT "Identification of ADAM 31: a protein expressed in Leydig cells and
RT specialized epithelia."
RL Endocrinology.141:2033-2042(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Bash G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.U., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni U., Mashima U., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald N., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: May be involved in sperm maturation and/or
CC fertilization. May also be involved in epithelia functions
CC associated with establishing and maintaining gradients of ions or
CC nutrients.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (Potential).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

```

```

CC -1- TISSUE SPECIFICITY: Highly expressed in Leydig cells. Expressed
CC also in cauda epididymidis, vas deferens, convoluted tubules,
CC kidney and the parietal cells of stomach. Not detected on
CC developing spermatocytes or mature sperm.
CC -1- DOMAIN: A tripeptide motif (VGE) within disintegrin-like domain
CC could be involved in the binding to egg integrin receptor and thus
CC could mediate sperm/egg binding.
CC -1- DOMAIN: The cysteine-rich domain encodes putative cell-fusion
CC peptides, which could be involved in sperm-egg fusion.
CC -1- PTM: Has no obvious cleavage site for furin endopeptidase,
CC suggesting that the proteolytic processing is regulated (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 disintegrin domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF251559; AAF74731.1; -
CC EMBL; AK014827; BAB29569.1; -
CC HSSP; P18619; 1FVL.
CC MEMOPS; M12.233; -.
CC MGD; MGI:1861229; Adam21.
CC InterPro; IPR005586; ADAM cysteine.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M2n_BS.
CC InterPro; IPR001590; peptidase_M12B.
CC InterPro; IPR002870; peptidase_M12B_N.
CC Pfam; PF00200; disintegrin; 1.
CC Pfam; PF01562; pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC PRINTS; PR00289; DISINTEGRIN.
CC PRODOM; PD000664; Disintegrin; 1.
CC SMART; SM00608; ACR; 1.
CC SMART; SM0050; DISIN; 1.
CC SMART; SM0181; EGF; 1.
CC PROSITE; PS50215; ADAM_MERO; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS50214; DISINTEGRIN_2; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01185; EGF_3; 1.
CC PROSITE; PS50026; EGF_3; 1.
CC HYDROLASE; PS00142; ZINC_PROTEASE; 1.
CC TRANSMEMBRANE; EGF-like domain.
CC SIGNAL 1 39
CC PROPEP 40 209
CC CHAIN 210 729
CC DOMAIN 210 685
CC TRANSMEM 686 706
CC DOMAIN 707 729
CC DOMAIN 210 402
CC DOMAIN 410 496
CC DOMAIN 497 639
CC DOMAIN 638 667
CC SITE 178 178
CC SITE 345 345
CC METAL 346 346
CC ACT SITE 349 349
CC METAL 355 355
CC DISULFID 320 397
CC DISULFID 360 382
CC DISULFID 468 481

```

```

FT DISULFID 638 649 BY SIMILARITY.
FT DISULFID 643 655 BY SIMILARITY.
FT DISULFID 657 666 BY SIMILARITY.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 729 AA; 80849 MW; E70B79BC46B4549 CRC64;

Query Match 37.0%; Score 42.5; DB 1; Length 729;
Best Local Similarity 36.4%; Pred. No. 56;
Matches 8; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

OY 1 HQCHOE-STXGRSRGRCGRGS 21
DB 531 HNCYKEINLQGNRFHCCTDGT 552

RESULT 13
YICG_CABEL STANDARD; PRT; 102 AA.
AC 08MNT7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Hypothetical UPF0287 protein C35D10.17 in chromosome III.
GN C35D10.17.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RA SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulcon L.;
RN Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
[2]
REVSTIONS.
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the UPF0287 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U21324; AAM2045.1.
DR WormPep; C35D10.17; C350625.
KW Hypothetical protein.
SQ SEQUENCE 102 AA; 11993 MW; C519494F4098A23D CRC64;

Query Match 36.5%; Score 42; DB 1; Length 102;
Best Local Similarity 40.0%; Pred. No. 9; 8;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 QCHQESTYGRSRGRC 16
DB 23 RCHSEKPIGKMGKGC 37

RESULT 14
LSM4_CABEL STANDARD; PRT; 123 AA.
AC Q19953;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.

```

```

GN F32A5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RA SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Paulley A.;
RN Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U20864; AAC4661.1.
DR PIR; T16234; T16234.
DR WormPep; F23A5.7; CE01277.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
KW RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 36.5%; Score 42; DB 1; Length 123;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 2 QCHQESTYGRSRGRCGRSG 20
DB 90 QCHQESTYGRSRGRCGRSG 108

RESULT 15
Y4CG_RHISN STANDARD; PRT; 305 AA.
AC P55389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Probable DNA-invertase Y4CG.
GN Y4CG.
OS Rhizobium sp. (strain NGR334).
OG Plasmid sym DNGR234.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RA SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: Belongs to the site-specific recombinase resolvase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>
CC or send an email to license@isb-sib.ch).

DR EMBL; AEC00068; AAB92422.1; -.

DR HSSP; P03012; 2KSL.
DR InterPro; IPR006118; Recombinase.

DR InterPro; IPR006119; resol_vase_N.
DR Pfam; PF00239; resol_vase: 1.

DR PROSITE: PS00397; RECOMBINASES_1; 1.
DR PROSITE: PS00398; RECOMBINASES_2; 1

KW Hypothetical protein; DNA recombination; DNA integration; DNA-binding; DNA invagination; Nucleoid

FT	ACT_SITE	23	23	TRANSIENT COVALENT LINKAGE TO DNA DURING
----	----------	----	----	------------------------------------------

STRAND CLEAVAGE AND REJOINING
(BY SIMILARITY).

SQ SEQUENCE 305 AA; 34277 MW; 48C03BD3AA49420F CRC64;

Query Match	36.5%	Score 42;	DB 1;	Length 305;
Best local Similarity	53.3%	Pred No	39;	

Matches: 8; Conservative: 1; Mismatches: 6; Indels: 0; Gaps: 0;

0Y 1 HOCHSTYGRGR 15

Db 42 HRIHQHSGASRAR 56

Search completed: September 28, 2004, 06:05:53
Took time: 0 00:00 0000

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 44.4062 Seconds

(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-10

Perfect score: 115
Sequence: 1 HQCHQESTXGRSGRGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	78.3	797	4	Q16824	Q16824 homo sapien
2	86	74.8	990	4	Q15206	Q15206 homo sapien
3	86	74.8	1218	4	Q05331	Q05331 homo sapien
4	79	68.7	798	4	Q95403	Q95403 homo sapien
5	79	68.7	1084	4	Q01212	Q01212 homo sapien
6	70	60.9	591	4	Q01720	Q01720 homo sapien
7	70	60.9	687	4	Q95402	Q95402 homo sapien
8	65	56.5	465	4	Q03838	Q03838 homo sapien
9	64	55.7	322	4	Q75370	Q75370 homo sapien
10	53.5	46.5	305	11	Q9CWM6	Q9CWM6 mus musculu
11	53	46.1	113	11	Q9CZ97	Q9CZ97 mus musculu
12	53	46.1	113	11	Q8BT22	Q8BT22 mus musculu
13	53	46.1	130	11	Q8CDR9	Q8CDR9 mus musculu
14	53	46.1	822	3	Q9P312	Q9P312 neurospora
15	53	46.1	947	12	Q8B489	Q8B489 dengue viru
16	52	45.2	1831	5	Q9W2E1	Q9W2E1 drosophila

17	52	45.2	1872	5	Q7YU83	Q7YU83 drosophila
18	50	43.5	214	16	Q8FIQ3	Q8FIQ3 corynebacte
19	49.5	43.0	305	11	Q9DCV4	Q9DCV4 mus musculu
20	49.5	43.0	778	11	Q7YU87	Q7YU87 mus musculu
21	49	42.6	678	10	Q8R200	Q8R200 oryza sativ
22	48	41.7	214	10	Q8SN94	Q8SN94 arabidopsis
23	48	41.7	254	10	Q8GMP2	Q8GMP2 arabidopsis
24	47.5	41.3	305	11	Q8C0Y9	Q8C0Y9 mus musculu
25	47.5	41.3	302	3	Q873D9	Q873D9 neurospora
26	47	40.9	128	4	Q9HBN1	Q9HBN1 homo sapien
27	47	40.9	157	10	Q84SD7	Q84SD7 oryza sativ
28	47	40.9	266	10	Q96410	Q96410 daucus caro
29	47	40.9	266	10	Q96411	Q96411 daucus caro
30	47	40.9	451	15	Q8B808	Q8B808 avian leuko
31	47	40.9	491	15	Q8U805	Q8U805 avian leuko
32	47	40.9	491	15	Q8U807	Q8U807 avian leuko
33	47	40.9	604	15	Q83132	Q83132 avian myelo
34	47	40.9	730	2	Q02301	Q02301 pseudomonas
35	47	40.9	1461	11	Q9TLP3	Q9TLP3 mus musculu
36	46.5	40.4	77	10	Q8H3P2	Q8H3P2 oryza sativ
37	46.5	40.4	338	16	Q9R1L1	Q9R1L1 deinococcus
38	46.5	40.4	443	10	Q9ARU6	Q9ARU6 oryza sativ
39	46	40.0	68	8	Q85WS9	Q85WS9 pinus korai
40	46	40.0	101	16	Q7VQD5	Q7VQD5 candidatus
41	46	40.0	170	10	Q94LT6	Q94LT6 oryza sativ
42	46	40.0	170	10	Q7YCH7	Q7YCH7 oryza sativ
43	46	40.0	268	10	Q96408	Q96408 daucus caro
44	46	40.0	644	16	Q89E32	Q89E32 bradyrhizob
45	46	40.0	849	10	Q9C7M2	Q9C7M2 arabidopsis

ALIGNMENTS

RESULT 1
ID Q16824 PRELIMINARY; PRT; 797 AA.

AC Q16824; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=1064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814]."
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60502; AAA63248.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF03516; Flaggrin; 4.
DR PRINTS; PR00487; FLAGGRIN.
FT NON_TER
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763SDA66B CRC64;

Query Match 78.3%; Score 90; DB 4; Length 797;
Best Local Similarity 81.0%; Pred. No. 8.8e-06;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HQCHQESTXGRSGRGRSGS 21
Db 425 HESHQESTXGRSGRGRSGS 445

RESULT 2

```

015206 ID 015206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=9125199; PubMed=2043621;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814 (1991).
DR EMBL: M60494; AAA63244.1;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR003303; Flaggrin.
DR Pfam: PF03516; Flaggrin; 6.
DR PRINTS: PR00487; FLAGGRIN.
FT NON TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 74.8%; Score 86; DB 4; Length 990;
Best Local Similarity 81.0%; Pred. No. 4.6e-05;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HOCHBESTXGSRGRCGRSGS 21
DB 225 HOCHBESTXGSRGRCGRSGS 245

RESULT 3
Q05331 ID 005331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FLAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Markov L.N., Chipev C.C., Gan S.Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-623 (1993).
RN [2]
RP FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
RN DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
RN TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
RN -1- PTM: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
RN HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
RN 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
RN DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
CC REPEATS.
DR EMBL: M96943; AAA36487.1; -.
DR PIR: A48118; A48118.
DR HSSP: P02593; 1CDM.
DR GO: GO:0005856; C:cytoskeleton; NAS.
DR GO: GO:0005509; F:calcium ion binding; NAS.
DR GO: GO:00030154; P:cell differentiation; NAS.
DR GO: GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro: IPR001751; CAP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Flaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Flaggrin; 6.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; FLAGGRIN.
DR PROSITE: PS00018; EF_HAND. 1.
DR PROSITE: PS00303; S100_CAP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA BIND 19 32 SITE I (BY SIMILARITY).
FT CA BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 1218 1218
SQ SEQUENCE 1218 AA; 13604 MW; EC195AD5285819C2 CRC64;

Query Match 74.8%; Score 86; DB 4; Length 1218;
Best Local Similarity 81.0%; Pred. No. 5.6e-05;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HOCHBESTXGSRGRCGRSGS 21
DB 447 HOCHBESTXGSRGRCGRSGS 467

RESULT 4
Q9H4U3 ID 09H4U3 PRELIMINARY; PRT; 798 AA.
AC Q9H4U3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE D14N1.1.2 (profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL356504; CAC13171.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR003303; Flaggrin.
DR Pfam: PF03516; Flaggrin; 4.
DR PRINTS: PR00487; FLAGGRIN.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 68.7%; Score 79; DB 4; Length 798;
Best Local Similarity 80.0%; Pred. No. 0.00047;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHBESTXGSRGRCGRSGS 21
DB 427 QCHBESTXGSRGRCGRSGS 446

RESULT 5
Q01212 ID 001212 PRELIMINARY; PRT; 1084 AA.

```



```

AC Q01212, Q03840;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published extratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL, M60501; AAA63243.1; JOINED.
DR EMBL, M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NMS.
DR GO; GO:0005198; F:structural molecule activity; NMS.
DR GO; GO:0007275; P:development; NMS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 68.7%; Score 79; DB 4; Length 1084;
Best Local Similarity 80.0%; Pred. No. 0.0063;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHOESTYGRSGRCGRSGS 21
DB 63 OSHOESTRGRSGRCGRSGS 82
ID Q01720 PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93054736; PubMed=1429717;
RA Freeland R.B., Haydock P.V., Fleckman P., Nurusukeiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene, genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
RT J. Biol. Chem. 267:23772-23781(1992).
DE -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DE DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
DE TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
DE -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
DE HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
DE 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
DE DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
DE PROBABLY INITIALLY CLEAVED
DE -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
DE WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
DE REPEATS.
DE EMBL, L01089; AAA60177.1; -.
DE EMBL, L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CABP; 1.
DR NON TER 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363BFEF07B74 CRC64;

Query Match 60.9%; Score 70; DB 4; Length 687;
Best Local Similarity 75.0%; Pred. No. 0.011;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 QCHOESTYGRSGRCGRSGS 21
DB 448 OSHOESTRGRSGRCGRSGS 467
ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE D014N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DE EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CABP; 1.
DR NON TER 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363BFEF07B74 CRC64;

```

```

QY      2 QCHOESTXGRSGRSGSGS 21
DB      448 QSHOESTRGRSGRSGSGS 467

RESULT 8
003838
ID      Q03838      PRELIMINARY;      PRT;      465 AA.
AC      Q03838;
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE      01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE      FLAGGRIN (PROFLAGGRIN) (Fragment).
GN      Homo sapiens (human).
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RX      MEDLINE=91064347; PubMed=2248957;
RA      Gan S.-Q., McBride O.W., Idler W.N., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
gene."
RL      Biochemistry 29:9432-9440 (1990).
RN      [2]
RP      REVISIONS.
RX      MEDLINE=91255199; PubMed=2043621;
RA      Gan S.-Q., McBride O.W., Idler W.N., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
gene."
RL      Biochemistry 30:5814-5814 (1991).
CC      -1- FUNCTION: FLAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
CC      PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC      FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC      -1- POLYMERISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC      WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
CC      REPEATS.
CC      -1- MISCELLANEOUS: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
CC      INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
CC      COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHVALIN
CC      GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC      PROTEOLYTICALLY CLEAVED.
DR      EMBL; M60499; AAA63246.1; -
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR003303; Flaggrin.
DR      Pfam; PF03516; Flaggrin; 3.
DR      PRINTS; PR00487; FLAGGRIN.
FT      NON_TER      1
FT      NON_TER      465
SQ      SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

Query Match      56.5%; Score 65; DB 4; Length 465;
Best Local Similarity 70.0%; Pred. No. 0.045;
Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 QCHOESTXGRSGRSGSGS 21
DB      226 QSHOESARSGRSGRSGSGS 245

RESULT 9
075370
ID      075370      PRELIMINARY;      PRT;      322 AA.
AC      075370;
DT      01-NOV-1998 (TReMBLrel. 08, Created)
DT      01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Epidermal filaggrin (Fragment).
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99101527; PubMed=9886436;
RA      Girbal-Neuhausser B., Durieux J.J., Arnaud M., Dalbon P., Sebbag M.,
RA      Vincent C., Simon M., Sersu T., Maeson-Bessiere C.,
RA      Jolivet-Reynaud C., Jolivet M., Serre G.;
RT      "The epitopes targeted by the rheumatoid arthritis-associated
RT      anti-flaggrin autoantibodies are posttranslationally generated on
RT      various sites of (pro)flaggrin by delamination of arginine residues."
RL      J. Immunol. 162:585-594 (1999).
DR      EMBL; AF043380; AAC23559.1; -
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR003303; Flaggrin.
DR      Pfam; PF03516; Flaggrin; 2.
DR      PRINTS; PR00487; FLAGGRIN.
FT      NON_TER      1
FT      NON_TER      322
SQ      SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match      55.7%; Score 64; DB 4; Length 322;
Best Local Similarity 68.4%; Pred. No. 0.045;
Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      2 QCHOESTXGRSGRSGSGS 20
DB      304 QSHOEFTRGSGRSGRSGS 322

RESULT 10
09CWS6
ID      09CWS6      PRELIMINARY;      PRT;      305 AA.
AC      09CWS6;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE      ES cells cDNA, RIKEN full-length enriched library, clone:241005016
DE      product:similar to CGI-90 protein.
GN      241005016RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RA      Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA      Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA      Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA      Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA      Kawai U., Kojima Y., Komno H., Kouda M., Koya S., Kurihara C.,
RA      Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA      Okazaki Y., Okido T., Owa C., Saito R., Saito R., Sakai C., Sakai K.,
RA      Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA      Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA      Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA      Muramatsu M., Hayashizaki Y.,
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium.
RA      the RIKEN Genome Exploration Research Group Phase I & II Team.
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs."
RL      Nature 420:563-573 (2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RX      MEDLINE=21085660; PubMed=11217851;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection."

```

```

RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of gap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
RT sequencing pipeline with 384 multichipillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
DR EMBL; AK010421; BAB26924.2; -
DR MGD; MGI:1913552; 241005016Rik.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
SQ SEQUENCE 305 AA; 35085 MW; 9EFF3012775F766 CRC64;

Query Match 46.1%; Score 53.5; DB 11; Length 305;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 12; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 3 HOCHSTXGSRGRCGRSG 20
DB 17 COQRTWG-SRGRCGRWG 33

RESULT 11
OQC297 PRELIMINARY; PRT; 113 AA.
AC OQC297;
DT 01-JUN-2001 (TRMBLrel. 17, Created)
DT 01-JUN-2001 (TRMBLrel. 17, Last sequence update)
DE 2810032G03Rik protein.
GN 2810032G03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Stadhil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni U., Mashima J., Mazzarelli U., Monchaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012851; BAB28512.1; -
DR MGD; MGI:1919919; 2810032G03Rik.
SQ SEQUENCE 113 AA; 13026 MW; A4DE33B5ED8F143 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 113;
Best Local Similarity 52.4%; Pred. No. 0.9;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 HOCHSTXGSRGRCGRSG 21
DB 3 HPCHESSVWRSMRRLGLGS 23

RESULT 12
OQB722 PRELIMINARY; PRT; 113 AA.
AC OQB722;
DT 01-MAR-2003 (TRMBLrel. 23, Created)
DT 01-MAR-2003 (TRMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN 2810032G03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028030; BAC25709.1; -
DR MGD; MGI:1919919; 2810032G03Rik.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 13010 MW; A4DE33B5ED8F143 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 113;
Best Local Similarity 52.4%; Pred. No. 0.9;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 HOCHSTXGSRGRCGRSG 21
DB 3 HPCHESSVWRSMRRLGLGS 23

RESULT 13
OQBDR9 PRELIMINARY; PRT; 130 AA.
AC OQBDR9;
DT 01-MAR-2003 (TRMBLrel. 23, Created)
DT 01-MAR-2003 (TRMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN 2810032G03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;

```

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/60; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=1246851;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK029678; BAC26560.1; -;
 DR MGD; MGI:1919919; 2810032G03Rik.
 KM Hypothetical protein.
 SQ SEQUENCE 130 AA; 14981 MW; AE5669C54B09D234 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 130;
 Best Local Similarity 52.4%; Pred. No. 1;
 Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

DQ 1 HCHCESTXGRSGRCRSGS 21
 20 HPCHESSVWRSMRRLGLGS 40

RESULT 14
 Q9P312 PRELIMINARY; PRT; 822 AA.
 AC Q9P312;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.
 GN B12FL10.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aligh V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AL390091; CAB98213.1; -;
 DR PIR; TS1049; TS1049.
 DR GO; GO:0003676; Functional acid binding; IEA.
 DR InterPro; IPR000504; RNA rec. mot.
 DR InterPro; IPR005120; Smg-4_DPF3.
 DR Pfam; PF00076; Rrm; 1.
 DR Pfam; PF03467; Smg4_DPF3; 1.
 DR SMART; SMO0360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PSS0030; RRM_RNP_1; 1.
 SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 46.1%; Score 53; DB 3; Length 822;
 Best Local Similarity 58.8%; Pred. No. 5.9;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DQ 5 QESTXGRSGRCRSGS 21
 414 RESASGTRGRGRGGT 430

RESULT 15
 Q9B489 PRELIMINARY; PRT; 947 AA.
 AC Q9B489;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NS5 protein (Fragment).

GN NS5.
 OS Dengue virus type 2.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=11060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Baleotti F.G., Koreli M.L., Figueiredo L.T.M.;
 RT "Detection of RNA polymerase motifs in the NS5 protein of Brazilian
 RT Flavivirus.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY167440; AA07510.1; -;
 DR GO; GO:0005524; F-ATP binding; IEA.
 DR GO; GO:0003658; F-RNA-directed RNA polymerase activity; IEA.
 DR InterPro; IPR00208; Flavi_NS5.
 DR InterPro; IPR002877; RrmU_Ftsd.
 DR Pfam; PF00972; Flavi_NS5; 1.
 DR Pfam; PF01728; FtsU; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 947 AA; 109981 MW; A644561B2864F8DC CRC64;

Query Match 46.1%; Score 53; DB 12; Length 947;
 Best Local Similarity 56.2%; Pred. No. 6.8;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

DQ 1 HCHCESTXGRSGRC 16
 849 HAHEEIQKGRGRRC 864

Search completed: September 28, 2004, 06:12:51
 Job time : 47.4052 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 59.8281 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-10

Perfect score: 115 HCHQBSTXGRSGRCGRSGS 21

Sequence: 1 HCHQBSTXGRSGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	113	98.3	21 2 AAW61520	AAW61520 Peptide X
2	113	98.3	22 2 AAE07235	AAE07235 IGP1546 P
3	86	74.8	1467 5 ABB97605	ABD97605 Novel hum
4	82	71.3	19 2 AAW61515	AAW61515 Peptide c
5	81	70.4	19 2 AAW61517	AAW61517 Peptide c
6	81	70.4	19 2 AAW61505	AAW61505 Peptide c
7	81	70.4	19 2 AAW61514	AAW61514 Peptide c
8	81	70.4	19 2 AAW61516	AAW61516 Peptide c
9	75	65.2	19 2 AAW61506	AAW61506 Peptide c
10	75	65.2	19 2 AAW61508	AAW61508 Peptide c
11	75	65.2	19 2 AAW61507	AAW61507 Peptide c
12	75	65.2	19 2 AAW61512	AAW61512 Peptide c
13	75	65.2	19 2 AAW61509	AAW61509 Peptide c
14	75	65.2	19 2 AAW61511	AAW61511 Peptide c
15	75	65.2	19 2 AAW61513	AAW61513 Peptide c
16	75	65.2	19 2 AAW61510	AAW61510 Peptide c
17	64	55.7	330 2 AAY22954	AAY22954 Human fl1
18	54.5	47.0	201 4 AAM79771	AAM79771 Human pro
19	54	47.0	70 5 ABP00440	ABP00440 Human ORF
20	54	47.0	330 2 AAY22956	AAY22956 Human fl1
21	54	47.0	330 2 AAY22955	AAY22955 Human fl1
22	54	47.0	330 2 AAY22957	AAY22957 Human fl1
23	52	45.2	424 4 ABB70938	ABB70938 Drosophila
24	50	43.5	477 6 ABC007142	ABC007142 Novel hum
25	50	43.5	1711 4 AAM79819	AAM79819 Human pro

26	50	43.5	1951 4 AAM78835	AAM78835 Human pro
27	48.5	42.2	488 2 AAR67757	AAR67757 Human fet
28	48.5	42.2	524 2 AAR67758	AAR67758 Human fet
29	48.5	42.2	670 2 AAR67759	AAR67759 Human fet
30	48.5	42.2	769 2 AAR75352	AAR75352 Human fet
31	48.5	42.2	961 7 ADE08457	ADE08457 Novel pro
32	48	41.7	214 5 ABB92582	ABB92582 Herbicida
33	48	41.7	641 4 ABG11945	ABG11945 Novel hum
34	47.5	41.3	854 7 ADD46732	ADD46732 Rat Prote
35	47.5	41.3	854 7 ADD46736	ADD46736 Rat Prote
36	47.5	41.3	854 7 ADB55432	ADB55432 Rat Prote
37	47	40.9	128 5 ABH04716	ABH04716 Human pP8
38	46.5	40.4	136 4 AAD28265	AAU28265 Novel hum
39	46	40.0	105 4 AAU61611	AAU61611 Propionib
40	46	40.0	105 6 ABM58130	ABM58130 Propionib
41	45	39.1	52 4 AAU56816	AAU56816 Propionib
42	45	39.1	52 6 ABM53355	ABM53355 Propionib
43	45	39.1	186 7 ADB64751	ADB64751 Human pro
44	45	39.1	211 5 AAU96183	AAU96183 Human sec
45	45	39.1	211 5 ABG64841	ABG64841 Human alb

ALIGNMENTS

RESULT 1
AAW61520
ID AAW61520 standard; peptide; 21 AA.
XX
AC AAW61520;
XX
DT 26-OCT-1998 (first entry)
XX
DB Peptide XI based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FT Key Location/Qualifiers
FT Modified-site 9 /note= "Citruilline"
FT XX
XX
XX W09822503-AA2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPEN.
XX Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Fig 1; 199p; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX
 SQ Sequence 21 AA;

Query Match 98.3%; Score 113; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 1 HCHQESTXGRSGRCGRSGS 21

RESULT 2
 AAB07235
 ID AAB07235 standard; peptide; 22 AA.
 AC AAB07235;
 XX
 XX 06-NOV-2001 (first entry)
 DT
 XX
 DE IGP1546 peptide for diagnosis and treatment of rheumatoid arthritis.
 XX
 KM Synthetic peptide; IGP1546; autoimmune antibody; rheumatoid arthritis;
 KM therapy; autoimmune disease; antirheumatic; systemic hyporesponsiveness;
 KM immunosuppressive; antiarthritic; cyclic.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Disulfide-bond 4..17
 FT Modified-site 10
 FT /note= "Citrulline"
 XX
 XX WO200146222-A2.
 XX
 PD 28-JUN-2001.
 XX
 PF 20-DEC-2000; 2000WO-EP013037.
 XX
 XX 21-DEC-1999; 99EP-00870280.
 PR 08-SEP-2000; 2000EP-00870195.
 XX
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Union A, Moereels H, Meheus L;
 PI
 XX WPI; 2001-49657/54.
 DR
 XX
 PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
 PT comprises citrulline residue between 2 cysteine residues and is
 PT specifically recognized by autoimmune antibodies from patients suffering
 PT from rheumatoid arthritis.
 XX
 XX Example 3; Page 32; 53pp; English.
 PS
 XX The present sequence is a synthetic peptide, IGP1546. The peptide of the
 CC invention comprises a citrulline residue between 2 cysteine residues and
 CC is specifically recognised by autoimmune antibodies from patients
 CC suffering from rheumatoid arthritis. The peptide comprises amino acids
 CC involved in side chain interactions which is essential for the formation
 CC of three-dimensional structure of the peptide. The peptide of the
 CC invention is useful as a medicament to treat autoimmune diseases,
 CC preferably rheumatoid arthritis. It is also useful for treating
 CC autoimmune diseases by increasing the size of antigen-immune complexes to
 CC improve clearance of the formed immune complexes and for the preparation
 CC of a medicament for oral or nasal administration to treat autoimmune
 CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
 CC to the peptide
 CC
 XX Sequence 22 AA;

Query Match 99.3%; Score 113; DB 4; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 2 HCHQESTXGRSGRCGRSGS 22

RESULT 3
 ABB97605
 ID ABB97605 standard; protein; 1467 AA.
 AC ABB97605;
 XX
 XX 27-JUN-2002 (first entry)
 DT
 XX
 DE Novel human protein SEQ ID NO: 873.
 XX
 KM Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
 KM antiferility; cerebroprotective; cyrostatic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 KM expressed sequence tag.
 XX
 OS Homo sapiens.
 XX
 PD WO200222660-A2.
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue J, Yang Y, Wehrman T, Dimanac RT;
 PI
 XX WPI; 2002-292408/33.
 DR N-PSDB; ABB32791.
 DR
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 PT
 XX Example 2; SEQ ID NO 873; 509pp; English.
 PS
 XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 CC
 XX Sequence 1467 AA;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 447 HCHQESTXGRSGRCGRSGS 467

RESULT 4
 AAM61515
 ID AAM61515 standard; peptide; 19 AA.
 XX

```

AC AAW61515;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfa, based on cDNA of a profilaggrin repeat.
DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX
XX Sequence 19 AA;
XX
XX
XX Query Match 71.3%; Score 82; DB 2; Length 19;
XX Best Local Similarity 88.9%; Pred. No. 3.4e-05;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX
XX 4 HOESTXGRSRGRCGRSGS 21
XX ||||| ||||| |||||
XX Db 2 HOESTAGRSRGRSGSGS 19
XX
XX
XX RESULT 5
XX AAW61517
XX ID AAW61517 standard; peptide; 19 AA.
XX
XX AAW61517;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfo, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KW solid phase synthesis; peptide amide; polyclonal antibody;
XX KW monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX W09822503-A2.
XX
XX

```

```

PD 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX
XX Sequence 19 AA;
XX
XX
XX Query Match 70.4%; Score 81; DB 2; Length 19;
XX Best Local Similarity 88.9%; Pred. No. 4.7e-05;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX
XX 4 HOESTXGRSRGRCGRSGS 21
XX ||||| ||||| |||||
XX Db 2 HOESTQGRSRGRCGRSGS 19
XX
XX
XX RESULT 6
XX AAW61505
XX ID AAW61505 standard; peptide; 19 AA.
XX
XX AAW61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KW solid phase synthesis; peptide amide; polyclonal antibody;
XX KW monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 7 /note="Citruilline"
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX

```

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

SQ Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 94.4%; Pred. No. 4.7e-05;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRCGRSGS 21

Db 2 HOESTXGRSRGRCGRSGS 19

RESULT 7

AAW61514

ID AAW61514 standard; peptide; 19 AA.

AC

XX AAW61514;

XX

DT 26-OCT-1998 (first entry)

XX

XX Peptide cf, based on cDNA of a proflilagrin repeat.

DE

XX

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KM solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX W09822503-A2.

PN

XX

PD 28-MAY-1998.

XX

XX

PF 14-NOV-1997; 97WO-NL000624.

XX

XX

PR 15-NOV-1996; 96NL-01004539.

XX

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

DR

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

SQ Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 88.9%; Pred. No. 4.7e-05;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRCGRSGS 21

Db 2 HOESTXGRSRGRCGRSGS 19

RESULT 8

AAW61516

ID AAW61516 standard; peptide; 19 AA.

XX

AC AAW61516;

XX

DT 26-OCT-1998 (first entry)

XX

XX Peptide cfE, based on cDNA of a proflilagrin repeat.

DE

XX

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KM solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX W09822503-A2.

PN

XX

PD 28-MAY-1998.

XX

XX

PF 14-NOV-1997; 97WO-NL000624.

XX

XX

PR 15-NOV-1996; 96NL-01004539.

XX

XX

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

DR

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

SQ Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 88.9%; Pred. No. 4.7e-05;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRCGRSGS 21

Db 2 HOESTXGRSRGRCGRSGS 19

RESULT 9

AAW61506


```

ID AAM61506 standard; peptide; 19 AA.
XX
XX AAM61506;
AC
XX 26-OCT-1998 (first entry)
DT
XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 9 /note="Citruiline"
FT
FT
XX W09822503-A2;
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WC-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX
XX Sequence 19 AA;
SQ
Query Match 65.2%; Score 75; DB 2; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 4 HOESTXGRSRGRCGRSGS 21
DB 2 HOESTXGRSRGRCGRSGS 19

```

```

XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 13 /note="Citruiline"
FT
FT
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WC-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX
XX Sequence 19 AA;
SQ
Query Match 65.2%; Score 75; DB 2; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 4 HOESTXGRSRGRCGRSGS 21
DB 2 HOESTRGRSRGXSGRSGS 19

```

```

RESULT 10
AAM61508
ID AAM61508 standard; peptide; 19 AA.
XX
XX AAM61508;
AC
XX 26-OCT-1998 (first entry)
DT
XX Peptide cfc4, based on cDNA of a profilaggrin repeat.
DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
XX
XX Sequence 19 AA;
SQ
Query Match 65.2%; Score 75; DB 2; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 4 HOESTXGRSRGRCGRSGS 21
DB 2 HOESTXGRSRGRCGRSGS 19

```

```

RESULT 11
AAM61507
ID AAM61507 standard; peptide; 19 AA.
XX
XX AAM61507;
AC
XX 26-OCT-1998 (first entry)
DT
XX Peptide cfc3, based on cDNA of a profilaggrin repeat.
DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 11 /note="Citruiline"
FT
FT
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX

```

```

XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX SQ Sequence 19 AA;

Query Match 65.2%; Score 75; DB 2; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 HOESTXGSRGRCGRSGS 21
DB 2 HOESTXGSRGRCGRSGS 19

RESULT 12
AAW61512
ID AAW61512 standard; peptide; 19 AA.
XX AC AAW61512;
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cfc8, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX OS Synthetic.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT Modified-site 7 /note= "Citruilline"
XX FT Modified-site 13 /note= "Citruilline"
XX FT Modified-site /note= "Citruilline"
XX PN WO9822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX PT

```

```

XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX SQ Sequence 19 AA;

Query Match 65.2%; Score 75; DB 2; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00036;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGSRGRCGRSGS 21
DB 2 HOESTXGSRGRCGRSGS 19

RESULT 13
AAW61509
ID AAW61509 standard; peptide; 19 AA.
XX AC AAW61509;
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclonal antibody.
XX OS Synthetic.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT Modified-site 16 /note= "Citruilline"
XX PN WO9822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX PT WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX
 SQ Sequence 19 AA:

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 83.3%; Pred. No. 0.00036;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRSGSGS 21
 |||||
 DB 2 HOESTXGRSRGRSGSGS 19

RESULT 14

AAW61511
 ID AAW61511 standard; peptide; 19 AA.

XX AAW61511;

DT 26-OCT-1998 (first entry)

DE Peptide cfc7, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key

FT Modified-site 7

FT Modified-site 11

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from

patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.

The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the

detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA:

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 88.9%; Pred. No. 0.00036;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRSGSGS 21
 |||||
 DB 2 HOESTXGRSRGRSGSGS 19

RESULT 15

AAW61513
 ID AAW61513 standard; peptide; 19 AA.

XX AAW61513;

DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key

FT Modified-site 7

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

FT Modified-site /note= "Citnulline"

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from

patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.

The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the

detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA:

QY 4 HOESTXGRSRGRSGSGS 21
 |||||
 DB 2 HOESTXGRSRGRSGSGS 19

Search completed: September 28, 2004, 06:24:22
Job time : 61.8281 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 / Search time 123.266 Seconds
(without alignments)

54.782 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HQCHQSTXGRSRKRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/ptcdat1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptcdat1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptcdat1/pubppaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptcdat1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptcdat1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptcdat1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptcdat1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptcdat1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptcdat1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptcdat1/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptcdat1/pubppaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptcdat1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptcdat1/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptcdat1/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptcdat1/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptcdat1/pubppaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/ptcdat1/pubppaa/US10_PUBCOMB.pep:*
- 18: /cgn2_6/ptcdat1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	98.3	21	US-09-308-150-10	Sequence 10, Appl
2	113	98.3	22	US-09-747-029A-22	Sequence 22, Appl
3	82	71.3	19	US-09-308-150-12	Sequence 12, Appl
4	81	70.4	19	US-09-308-150-11	Sequence 11, Appl
5	81	70.4	19	US-09-308-150-11	Sequence 11, Appl
6	81	70.4	19	US-09-308-150-13	Sequence 13, Appl
7	81	70.4	19	US-09-308-150-14	Sequence 14, Appl
8	75	65.2	19	US-09-308-150-2	Sequence 2, Appl
9	75	65.2	19	US-09-308-150-3	Sequence 3, Appl
10	75	65.2	19	US-09-308-150-4	Sequence 4, Appl
11	75	65.2	19	US-09-308-150-5	Sequence 5, Appl
12	75	65.2	19	US-09-308-150-6	Sequence 6, Appl
13	75	65.2	19	US-09-308-150-7	Sequence 7, Appl
14	75	65.2	19	US-09-308-150-8	Sequence 8, Appl
15	75	65.2	19	US-09-308-150-9	Sequence 9, Appl

16 52 45.2 172 16 US-10-437-963-175440 Sequence 175440,
17 50 43.5 206 16 US-10-437-963-110381 Sequence 110381,
18 50 43.5 477 15 US-10-161-927-62 Sequence 62, Appl
19 49 42.6 103 16 US-10-437-963-122475 Sequence 122475,
20 49 42.6 110 16 US-10-437-963-104204 Sequence 104204,
21 49 42.6 150 16 US-10-437-963-172443 Sequence 172443,
22 48.5 42.2 88 16 US-10-437-963-131735 Sequence 131735,
23 48.5 42.2 160 12 US-10-425-114-66342 Sequence 66342, A
24 48.5 42.2 526 15 US-10-369-993-10496 Sequence 10496, A
25 48 41.7 1087 16 US-10-437-963-123162 Sequence 123162,
26 48 41.7 1087 9 US-09-918-909-24 Sequence 24, Appl
27 48 41.7 1087 16 US-10-641-991-24 Sequence 24, Appl
28 47.5 41.3 176 16 US-10-437-963-159623 Sequence 159623,
29 47 40.9 201 16 US-10-437-963-174936 Sequence 174936,
30 47 40.9 218 12 US-10-425-114-68534 Sequence 48534, A
31 47 40.9 363 16 US-10-437-963-151018 Sequence 151018,
32 47 40.9 384 12 US-10-425-114-72136 Sequence 72136, A
33 47 40.9 384 12 US-10-425-114-72137 Sequence 72137, A
34 46.5 40.4 56 16 US-10-437-963-155194 Sequence 155194,
35 46.5 40.4 136 12 US-10-221-278-622 Sequence 622, App
36 46.5 40.4 136 15 US-10-231-172-622 Sequence 622, App
37 46.5 40.4 456 16 US-10-437-963-155551 Sequence 155551,
38 46 40.0 53 16 US-10-767-701-57837 Sequence 57837, A
39 46 40.0 108 16 US-10-437-963-147392 Sequence 147392,
40 46 40.0 506 16 US-10-437-963-185974 Sequence 185974,
41 46 40.0 807 12 US-10-425-114-43305 Sequence 43305, A
42 46 40.0 807 12 US-10-425-114-45403 Sequence 44503, A
43 46 40.0 840 16 US-10-437-963-151528 Sequence 151528,
44 45 39.1 42 14 US-10-029-386-31444 Sequence 31444, A
45 45 39.1 164 16 US-10-437-963-112419 Sequence 112419,

ALIGNMENTS

RESULT 1
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Vennocij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

Query Match 98.3%; Score 113; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;

	Matches	21; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1	HCHCHSTYGRSRGRGRGSS	21		
Db	1	HCHCHSTYGRSRGRGRGSS	21		

```

RESULT 2
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Am
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydie
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.003INPUS00 INNS:031
CURRENT APPLICATION NUMBER: US/09/747,029A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptides
NAME/KEY: MOD RES
LOCATION: (10)..(10)
OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

```

Query Match	98.3%	Score 113;	DB 9;	Length 22;
Best Local Similarity	100.0%	Pred. No. 1.3e-08;		
Matches	21;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

```

RESULT 3
US-09-308-150-12
; Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raets, Jozef Maria Hendrik
APPLICANT: Hoer, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO. 12
LENGTH: 19
TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilin
DS-09-308-150-12

```

```

Query Match      71.3%; Score 82; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 4 HOESTYGRGRGRCGRSGS 21
    |||||
Db 2 HOESTAGRGGRGRCGRSGS 19

```

```

RESULT 4
US-09-308-150-1
: Sequence 1. Application US/09308150
: Patent No US20020137092A1
:
GENERAL INFORMATION:
: APPLICANT: Van Vennrooij, Waltherus Jacobus Wilhelmus
: APPLICANT: Schellekens, Gerardus Antonius
: APPLICANT: Raats, Jozef Maria Hendrik
: APPLICANT: Hoeft, Rene Michael Antonius
: APPLICANT: Stichting Scheikundig Onderzoek Nederland
: APPLICANT: Stichting voor de Technische Wetenschappen
:
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
:
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
:
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
:
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
:
CURRENT APPLICATION NUMBER: US/09/308,150
:
CURRENT FILING DATE: 1999-09-30
:
PRIOR APPLICATION NUMBER: PCT/NL97/00624
:
PRIOR FILING DATE: 1997-11-14
:
PRIOR APPLICATION NUMBER: NL 1004539
:
PRIOR FILING DATE: 1996-11-15
:
NUMBER OF SEQ ID NOS: 16
:
SOFTWARE: PatentIn Ver. 2.1
:
SEQ ID NO 1
:
LENGTH: 19
:
TYPE: PRT
:
ORGANISM: Artificial Sequence
:
FEATURE:
:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
:
OTHER INFORMATION: known cDNA sequences of human profilaggrin
:
OTHER INFORMATION: Xaa is citrulline
:
US-09-308-150-1

```

Query Match	70.4%;	Score 81;	DB 9;	Length 19;
Best Local Similarity	94.4%;	Pred. No. 0.00026;		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```

RESULT 5
US-09-308-150-11
; Sequence 11, Application US/09308150
; Patent No. US20020137092A1
;
; GENERAL INFORMATION:
; APPLICANT: Van Veenrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```

FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-11

Query Match 70.4%; Score 81; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00026;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTGRSRGRSGSGS 21
DB 2 HOESTGRSRGRSGSGS 19

RESULT 6
US-09-308-150-13

Sequence 13, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoer, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-13

Query Match 70.4%; Score 81; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00026;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTGRSRGRSGSGS 21
DB 2 HOESTGRSRGRSGSGS 19

RESULT 7
US-09-308-150-14
Sequence 14, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoer, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-14

Query Match 70.4%; Score 81; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00026;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTGRSRGRSGSGS 21
DB 2 HOESTGRSRGRSGSGS 19

RESULT 8
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoer, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-2

Query Match 65.2%; Score 75; DB 9; Length 19;

Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21
Db 2 HOESTRGRSGRSGRSGS 19

RESULT 9

US-09-308-150-3

Sequence 3, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-3

Query Match 65.2%; Score 75; DB 9; Length 19;

Best Local Similarity 83.3%; Pred. No. 0.0017; 3; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTRGRSGRSGRSGS 19

RESULT 10

US-09-308-150-4

Sequence 4, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTRGRSGRSGRSGS 19

RESULT 11

US-09-308-150-5

Sequence 5, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-5

Query Match 65.2%; Score 75; DB 9; Length 19;

Best Local Similarity 83.3%; Pred. No. 0.0017; 3; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTRGRSGRSGRSGS 19

RESULT 12

US-09-308-150-6

Sequence 6, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRGRSGS 21
|||||
Db 2 HOESTXGRSGRGRSGS 19

RESULT 13

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRGRSGS 21

Db 2 HOESTXGRSGRGRSGS 19
|||||

RESULT 14

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRGRSGS 21
|||||
Db 2 HOESTXGRSGRGRSGS 19

RESULT 15

US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. NO. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGSRGRGCGSGS 21
|||
Db 2 HOESTXGSRGRGCGSGS 19

Search completed: September 28, 2004, 07:28:52
Job time : 123.266 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 15.9688 Seconds

(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HQCHQESTXGRSGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.5	43.0	234	4	US-09-252-991A-17302 Sequence 17302, A
2	49.5	42.6	211	4	US-09-252-991A-24706 Sequence 24706, A
3	48.5	42.2	236	4	US-09-252-991A-25980 Sequence 25980, A
4	48.5	42.2	488	1	US-08-243-542-1 Sequence 1, Appli
5	48.5	42.2	488	1	US-08-477-407-1 Sequence 1, Appli
6	48.5	42.2	488	1	US-08-484-355-1 Sequence 1, Appli
7	48.5	42.2	524	1	US-08-243-542-2 Sequence 2, Appli
8	48.5	42.2	524	1	US-08-477-407-2 Sequence 2, Appli
9	48.5	42.2	524	1	US-08-484-355-2 Sequence 2, Appli
10	48.5	42.2	670	1	US-08-243-542-3 Sequence 3, Appli
11	48.5	42.2	670	1	US-08-477-407-3 Sequence 3, Appli
12	48.5	42.2	670	1	US-08-484-355-3 Sequence 3, Appli
13	48.5	42.2	769	1	US-08-243-542-4 Sequence 4, Appli
14	48.5	42.2	769	1	US-08-477-407-4 Sequence 4, Appli
15	48.5	42.2	769	1	US-08-484-355-4 Sequence 4, Appli
16	46.0	40.0	255	4	US-09-252-991A-19598 Sequence 19598, A
17	45.0	39.1	105	4	US-09-252-991A-12393 Sequence 12393, A
18	45.0	39.1	231	4	US-09-252-991A-17218 Sequence 17218, A
19	45.0	39.1	267	4	US-09-252-991A-18758 Sequence 18758, A
20	45.0	39.1	352	4	US-09-252-991A-18590 Sequence 18590, A
21	45.0	39.1	363	4	US-09-252-991A-18512 Sequence 18512, A
22	45.0	39.1	406	4	US-09-252-991A-19857 Sequence 19857, A
23	45.0	39.1	421	4	US-09-252-991A-23366 Sequence 23366, A
24	45.0	39.1	1151	4	US-09-252-991A-21328 Sequence 21328, A
25	44.5	38.7	83	4	US-09-716-129-151 Sequence 151, App
26	44.5	38.7	696	3	US-09-351-414-2 Sequence 2, Appli
27	44.5	38.7	834	4	US-09-252-991A-17616 Sequence 17616, A

28	44.5	38.7	923	4	US-09-252-991A-27148 Sequence 27148, A
29	44.5	38.7	957	4	US-09-252-991A-20408 Sequence 20408, A
30	43.5	37.8	62	4	US-09-252-991A-32266 Sequence 32266, A
31	43.5	37.8	212	4	US-09-252-991A-29831 Sequence 29831, A
32	43.5	37.8	292	4	US-09-199-637A-419 Sequence 419, App
33	43.0	37.4	178	4	US-09-252-991A-23975 Sequence 23975, A
34	43.0	37.4	203	4	US-09-252-991A-26436 Sequence 26436, A
35	43.0	37.4	225	4	US-09-252-991A-17754 Sequence 17754, A
36	43.0	37.4	865	4	US-09-252-991A-18683 Sequence 18683, A
37	43.0	37.4	1091	4	US-09-252-991A-25637 Sequence 25637, A
38	43.0	37.4	1253	1	US-08-252-966B-12 Sequence 12, Appli
39	42.5	37.0	591	4	US-09-252-991A-28760 Sequence 28760, A
40	42.0	36.5	175	4	US-09-252-991A-19711 Sequence 19711, A
41	42.0	36.5	404	4	US-09-489-039A-11606 Sequence 11606, A
42	42.0	36.5	405	4	US-09-252-991A-20011 Sequence 20011, A
43	42.0	36.5	411	4	US-09-252-991A-22445 Sequence 22445, A
44	42.0	36.5	483	4	US-09-252-991A-19015 Sequence 19015, A
45	42.0	36.5	518	4	US-09-252-991A-25967 Sequence 25967, A

ALIGNMENTS

```
RESULT 1
US-09-252-991A-17302
; Sequence 17302, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17302
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17302

Query Match          43.0%; Score 49.5; DB 4; Length 234;
Best Local Similarity 45.5%; Pred. No. 5.4;
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY      1 HQCHQESTXG-----RSRRCG 17
Db      76 HECRDPTRGHDGSPRAPRCG 97

RESULT 2
US-09-252-991A-24706
; Sequence 24706, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24706
; LENGTH: 211
; TYPE: PRT
```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24706

Query Match
Best Local Similarity 42.6%; Score 49; DB 4; Length 211;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 QESTXGRSGRCGRSG 20
DB 42 RSSAGRSRRRCARAG 57

RESULT 3
US-09-252-991A-25980
Sequence 25980, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107136.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25980
LENGTH: 236
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25980

Query Match
Best Local Similarity 42.2%; Score 48.5; DB 4; Length 236;
Matches 11; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 4 HOESTXGRS-----RCGRCS 19
DB 78 HRRHROGRSRDOGRRCGRCS 98

RESULT 4
US-08-243-542-1
Sequence 1, Application US/08243542
Patent No. 5552526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-1

Query Match
Best Local Similarity 42.2%; Score 48.5; DB 1; Length 488;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHQESTXGRSGRCGRSGS 21
DB 461 CYEKLNVGTERGSCGRKGS 480

RESULT 5
US-08-477-407-1
Sequence 1, Application US/08477407
Patent No. 5631351
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-477-407-1

Query Match 42.2%; Score 48.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO6-STXGRSGRCGRSGS 21
Db 461 CYEKLNVETGSGCRKGS 480

RESULT 6
US-08-484-355-1
Sequence 1, Application US/08484355
Patent No. 5705341
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAS
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Wordperfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:

TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-484-355-1

Query Match 42.2%; Score 48.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO6-STXGRSGRCGRSGS 21
Db 461 CYEKLNVETGSGCRKGS 480

RESULT 7
US-08-243-542-2
Sequence 2, Application US/08243542
Patent No. 5532526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAS
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Wordperfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 524 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSRGCRGSGS 21
DB 468 CYEKLNVGTERGSGCRKGS 487

RESULT 8
US-08-477-407-2
Sequence 2, Application US/08477407

PATENT No. 5631351
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994

APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Teriyence F. Chapman
REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 524 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-477-407-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSRGCRGSGS 21
DB 468 CYEKLNVGTERGSGCRKGS 487

RESULT 9
US-08-484-355-2
Sequence 2, Application US/08484355
PATENT No. 5705341

GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994

APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Teriyence F. Chapman
REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 524 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-484-355-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSRGCRGSGS 21
DB 468 CYEKLNVGTERGSGCRKGS 487

Db 468 CYEKLINVEGTERGSCGRKGS 487

RESULT 10

US-08-243-542-3
; Sequence 3, Application US/08243542
; Patent No. 5552526
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: : EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Wordperfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,542
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 670 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
US-08-243-542-3

Query Match 42.2%; Score 48.5; DB 1; Length 670;
Best Local Similarity 45.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHOE-STXGRSGRGRSGS 21
Db 468 CYEKLINVEGTERGSCGRKGS 487

RESULT 11
US-08-477-407-3
; Sequence 3, Application US/08477407
; Patent No. 5631351
; GENERAL INFORMATION:

; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: : EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Wordperfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,407
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 670 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
US-08-477-407-3

Query Match 42.2%; Score 48.5; DB 1; Length 670;
Best Local Similarity 45.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHOE-STXGRSGRGRSGS 21
Db 468 CYEKLINVEGTERGSCGRKGS 487

RESULT 12
US-08-484-355-3
; Sequence 3, Application US/08484355
; Patent No. 5705341
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: : EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terrence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 670 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-484-355-3
Query Match 42.2%; Score 48.5; DB 1; Length 670;
Best Local Similarity 45.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 3 CHOE-STXGRSRRCGRSGS 21
DB 468 CYEKLNVGTERGSCGRKGS 487
RESULT 13
US-08-243-542-4
Sequence 4, Application US/08243542
Patent No. 5532526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terrence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-4
Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 3 CHOE-STXGRSRRCGRSGS 21
DB 567 CYEKLNVGTERGSCGRKGS 586
RESULT 14
US-08-477-407-4
Sequence 4, Application US/08477407
Patent No. 5631351
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995

Page 7

```

      APPLICATION NUMBER: JP 5-257455
      FILING DATE: 22 SEPTEMBER 1993
      APPLICATION NUMBER: JP 6-49904
      FILING DATE: 23 FEBRUARY 1994
      APPLICATION NUMBER: JP 6-73328
      FILING DATE: 12 APRIL 1994
      APPLICATION NUMBER: JP 6-84470
      FILING DATE: 22 APRIL 1994
      ATTORNEY/AGENT INFORMATION:
      NAME: Terryence F. Chapman
      REGISTRATION NUMBER: 32 549
      REFERENCE/DOCKET NUMBER: Futuya Case 1313
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (616) 381-1156
      TELEFAX: (616) 381-5465
      INFORMATION FOR SEQ ID NO: 4:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 769 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      ORIGINAL SOURCE:
      ORGANISM: Homo sapiens
      IMMEDIATE SOURCE:
      LIBRARY: human fetal brain cDNA library
      US-08-484-355-4
      Query Match 42.2%; Score 48.5; DB 1; Length 769;
      Best Local Similarity 45.0%; Pred. No. 25;
      Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1.
      QY 3 CHOE-STYGRSRGRCGRSGS 21
      :::: |||||
      Db 567 CYEKLNVGTERSGRCGRGS 586

```

Blank sheet